

What to make of nuclear winter

The International Council of Scientific Unions (ICSU) report on nuclear winter, the supposed aftermath of nuclear warfare, is the best on this subject so far. But the policy implications are not clear.

So do we know what are the policy effects of nuclear war? The question is naturally prompted by the appearance last week of the monumental report by the working party called the Scientific Committee on Problems of the Environment (SCOPE) set up three years ago by the International Council of Scientific Unions (ICSU). There are two senses in which the answer can only be hesitant assent, neither of which detracts from the value of what the committee has done. First, like all previous studies in this field going back to the then invaluable document published by the government of India in 1957, and to nobody's surprise, such documents inevitably draw attention to problems whose solutions are not to hand, but whose reality had not previously been apparent. Second, there is an uncertainty of a different kind arising because nobody can be really sure how a nuclear war would be conducted by the military. The usual starting assumption is that the amount of nuclear explosives used would be a substantial fraction of the arsenals of the major powers, which is fair enough. The conclusions are usually (as in the SCOPE study) that a nuclear war would be a near-catastrophe, from which some go on to conclude that the steps that some governments take to limit anticipated damage (such as civil defence) are quite beside the point. But there are all too many reasons for believing that nuclear wars falling short of full-scale catastrophe are the most likely — an unhappy circumstance because it can lead to the concept of a "winnable" nuclear conflict.

None of this diminishes the value of the SCOPE study, which is best regarded as the latest contribution to a sequence of studies still far from complete. (Last week's report, like many of its predecessors, is also an agenda for research.) But this study, to SCOPE's credit, has been carried out in an open fashion. There have been opportunities for those who have had special knowledge to speak up at workshops and seminars, the functions of which have been advertised in advance. As a model for the conduct of inquiries in fields in which the uncertainties are so huge, SCOPE seems to have broken useful new ground. That the actual release of its report should have fallen into the hands of publicity managers more free with press releases than with copies of the full text (not yet available) is probably best explained by the need in the United States to shout even to be heard.

Fallout

It is nevertheless a pity that the SCOPE report will be distinguished from its predecessors chiefly because it provides an assessment of the proposition that nuclear winter may follow nuclear war. The other sections of the document, for example that dealing with the estimation of the effects of radioactive fallout after the explosion of several megatons of bombs near the surface of the Earth, is in itself an important argument that should not be overlooked. (The conclusions are more gloomy than some previous calculations.) Yet nuclear winter needs pride of place (and gets it) because it is a novel idea only recently imported into the discussion of what will happen to those unfortunate enough to survive the immediate effects of a nuclear war. The case deserves a hearing, which it will get at least from those who may hitherto have thought that geography would insulate them from the most damaging effects of nuclear war.

What SCOPE says about nuclear winter does not differ qual-

itatively, or in other significant ways, from the earlier accounts of nuclear winter (see Turco *et al.* *Science* **222**, 1283; 1983). Smoke from fires ignited by nuclear explosions will be lifted high into the troposphere, the Sun will be obscured, the surface of the Earth will be cooled, photosynthesis will be stopped in many places and animals (including people) will die. The effects will be global, extending beyond the frontiers of the states directly involved with the war (which is why states such as India are alarmed). Nuclear winters may also be seasonal, with smoke from summer wars spreading more quickly. Given the uncertainties of calculating nuclear winters, it is neither here nor there that SCOPE's estimates of the severity, duration and distribution of the effects of nuclear winter are on the cheerful side of those of Turco *et al.* The differences are small, and not important. For the time being, the nuclear winter must firmly be listed among the consequences of substantial nuclear wars, remembering that time (but, it is hoped, not experience) will show that even present fears are immoderate. SCOPE's antecedents require no less.

Uncertainties

It is also important that the uncertainties should not be buried by last week's razzamatazz. The SCOPE report acknowledges that the severity of a nuclear winter will be a function of the quantity of smoke remaining in the atmosphere, which is most probably not a linear function of the amount injected but rather an S-shaped function rising to a plateau. The quantities are huge, measured in tens of millions of tons of elemental carbon in the form of micrometre-sized particles. Fifty million tons will be more than enough to make a serious nuclear winter. (The upper range of the SCOPE estimates is some 30 million tons of smoke in the atmosphere.) Twice as much will not make things very much worse. The crucial question for the theory of the nuclear winter is the proportion of the smoke that can be expected to stay in the atmosphere. SCOPE properly lists the now familiar uncertainties, the effects of rain-clouds generated by rising plumes of heat, coagulation and so on. The report also allows that little is so far known of the atmospheric processes on the mesoscale (between that of a smoke plume and that of a continent) which may have an important influence on the removal of smoke, but from now on it will probably be wise to follow SCOPE's conclusion that nuclear winter is a probable consequence of nuclear war.

So what? The SCOPE report, for all its bulk, has steered clear of taking up the implications of its conclusions. But Sir Frederick Warner, chairman of the steering committee warned that the policy implications cannot be ignored. But what are they? For forty years, many people and most governments have behaved as if they believed nuclear war to be such a potential catastrophe that international relations should be moulded so as to avoid it. These good intentions have not been conspicuously successful. The possibility now that the effects of nuclear war will be spread beyond the territory of the combatants gives non-nuclear countries an even greater interest in persuading the nuclear powers to arms control. But the countries concerned are already crying with frustration at their inability to persuade their bigger brothers down the path of mutual restraint. The certainty that billions might die in neutral populations is no more likely to



affect the strategic relationship between the superpowers than the certainty, acknowledged for decades, that some tens of million would be put in hazard from radioactive fallout. The reasons of course are not technical but political. The SCOPE committee might have taken its courage in both hands and said that the two activities cannot be separated.

SCOPE seems to argue instead that it is for technical committees to stick to the production of technical reports and for others, called politicians, to interpret them. The obvious danger is that technical arguments and conclusions will be misinterpreted. Sometimes, they may even seem to be used mischievously. But how can that be done when the conclusion of a technical report is as factual as the statement that large-scale nuclear war may have consequences even more damaging than previously thought? The United States government is right to say that a democracy as open as it is itself, is more vulnerable than its adversaries to public fears of the results of present policies. Its claim on the sympathy even of its friends is however diminished as its own policy on arms control becomes to seem more like a means of keeping things the way they are. The chances are high that the SCOPE report will be used in ways like these to persuade governments to make concessions leading to action that may be unwise. SCOPE might have acknowledged that the danger exists.

What else might have been said? With the review conference of the Non-Proliferation Treaty coming to what seems (against the odds) to be an unexpectedly amicable end, but with several months of arms control negotiations, summit meetings and diplomatic troubles lying ahead, SCOPE might have acknowledged the subtlety of the problems facing governments. It is a sensible starting point to take 600 megatons as the assumed quantity of nuclear explosives likely to be used in a major nuclear war, but governments may also fight lesser wars, where one or two nuclear bombs may be used in anger before good sense gets the better of the combatants, and they stop. (That is where civil defence comes in.) And while countries that calculate that even nuclear war would leave them immune from direct attack, but which may find nuclear winter a novel threat, should be reminded that they have as much to lose from the general collapse of international order. That is a point that SCOPE does make.

What does this mean for research? The SCOPE report follows its predecessors in listing topics that call for urgent study. Many of these are problems that need more effort on other grounds in any case. Mesoscale processes in the atmosphere are a kind of missing link in climatology and meteorology. The effect of clouds (real clouds, not average cloudiness) on the insolation of the surface of the Earth is crucial to a proper understanding of the carbon dioxide problem, yet understanding is far from complete. These are problems that need most urgently to be taken up. The prospect that they may conspire to produce a nuclear winter after a nuclear war should be noted as yet another proof that nuclear wars are to be avoided, and such funds as there are for more research should be put where the scientific problems are. In a decade or so, SCOPE would probably be able to produce an even better document. The trouble is that the amplifiers of SCOPE's conclusions are already being tuned for action. When the United Nations secretariat wrote to *Nature* a few days ago requesting permission to reprint extracts from articles on nuclear winter, it became plain that its claims on others' copy-right were curiously one-sided. The UN document is due to be published next year. □

Bees in South-East Asia

The Wall Street Journal's attack on Professor M. Meselson's is unwarranted and unsubstantiated.

THE *Wall Street Journal*, widely acclaimed as one of the best newspapers in the world, seems to have a quirky streak that it should exorcise. Although the journal's title suggests a paper designed to bring comfort to those whose interests extend no further than that day's deal, in reality the journal is almost

always the first with the telling confession of the latest bank to have gone to the wall. More than that, the newspaper tells its readers perceptively of the best plays wherever they are put on in the United States, and it is perhaps most of all distinguished by its attempts, mostly successful, to account for economic trends by anecdotal accounts of human behaviour — often these days a painful interview with a farmer giving up the land — and the technical information (how may ECU to the dollar for example) is superb. The most obvious curiosity is that the thoroughly objective but adventurous body of the paper sandwiches each day two editorial pages whose temper is that of the unreconstructed conservative. This does not mean that the journal supports the present government of the United States: that is more often denounced for backsliding from some public position or election promise. President Reagan often comes through as a kind of "pinko". Curiously, for a newspaper so technically advanced that it can replicate itself at more places in the United States than any other, the *Wall Street Journal* pays very little attention to science, which might be thought to be its bread and butter for the long term.

Except occasionally. Two years ago, the journal startled its readers with a sensational series of articles spelling out the way in which the Soviet Union was supposed to be using biotechnology as a way of developing new kinds of biological warfare. The evil genius behind these schemes was supposed to be Academician Ovchinnikov, vice-president of the Soviet Academy of Sciences, who inconsistently more often gives the impression that his cup of happiness would be full if only he were to succeed Academician Aleksandrov as president. Luckily, largely because of the internal inconsistencies, nobody of much importance seems to have taken the biological warfare scare seriously. The author of the series of *Wall Street Journal* articles was one William Kucewicz, described as a member of the editorial board of the newspaper.

Mr Kucewicz last week burst out again, with an attack on Professor Matthew Meselson's claim that many of the reports of "yellow rain" in South-East Asia could be explained naturally as the excrement of bees. Meselson and some of his associates have found material such as this in South-East Asia, although not where the material is said to have been used malevolently, to damage people (see *Nature* 309, 205; 1985).

There is no obvious reason why people wishing to harm populations with mycotoxins derived from fungi should choose to distribute them in intimate association with pollen particles (which occur in some samples from the Vietnamese battlefields, and in Laos, where the accusations first surfaced). Meselson seems never to have claimed that the absence of mycotoxins in particles of yellow rain, whatever their origin, would be a proof that the Soviet Union could be defended from the charge of using the countries bordering Vietnam as a proving ground for new weapons.

Mr Kucewicz's attack on this position is curiously blunt. The "bee faeces" theory is discredited, he claims, on the basis of Meselson's confession (when asked) that there were no traces of mycotoxin in the samples of bee excrement recovered away from the Vietnam battle areas. What, of course, this discovery implies is one of the possibilities suggested a few years ago, that toxic yellow rain might be made from pollen that had been naturally contaminated, perhaps by fungal invasion. To claim that the absence of mycotoxins in yellow rain from one place proves that there are mycotoxins in yellow rain from another place, and that they must have got there by design, presumably Soviet, makes no sense that can easily be unravelled. The truth, of course, is that pollen, airborne or from the guts of bees, was never a plausible vehicle for a toxin. To confess that yellow rain all comes from bees would not in itself prove that toxins have never been used in South-East Asia. Other vehicles might have been used. What the bee theory does however do is to discredit the investigations on which allegations of the Soviet use of biological weapons have been based. The simplest solution would be for the State Department to publish the evidence on which the allegations are based. □

Nuclear winter

International committee echoes gloomy forecasts

Washington

THE study of the environmental consequences of nuclear war made public last week by the Scientific Committee on Problems of the Environment (SCOPE) is the most ambitious attempt yet to arrive at a comprehensive view of the problem. The report will be published in two volumes, and runs to 800 pages (see also p.192). The immediate effects of nuclear explosions and of radioactive fallout are considered, but it is inevitable that most attention will be given to what the study says about nuclear winter and the probable biological consequences for both natural and agricultural ecosystems that would follow. The modelling in the study "extends very greatly" previous work, according to the chairman of the study's steering committee, Sir Frederick Warner of the University of Essex in England.

The principal authors of the report acknowledge the many uncertainties surrounding estimates of the amount of smoke likely to be released from mass fires and its residence time in the atmosphere, two key factors for the nuclear winter principle. But the study found that all model simulations considered suggested "a strong potential for large-scale weather disruptions as a result of smoke injected by post-nuclear fires".

Although simplifications of the models and uncertainties about some physical processes may affect the details of predictions, they "probably do not affect the general character of the calculated atmospheric response". Tentative estimates indicate that for a nuclear exchange in summer months in the Northern Hemisphere with 6,500 megatonne (MT) equivalent of weapons exploded, a temperature drop of between 15 and 35° C in the few weeks after an attack is plausible for continental interiors in the northern midlatitudes, with the day-to-day values depending on the presence or otherwise of patchy dense smoke. (But SCOPE does not believe that thresholds in the relation between depth of smoke and the amount of radiation reaching the ground imply that there is an "acceptable" amount of smoke below which cooling effects do not occur.)

Other places in the Northern Hemisphere, especially coastal areas and islands, might experience much smaller temperature drops of between 0 and 5° C, because of the effects of convection near the ocean/land boundary. But long-term effects might last for years and would slowly spill over into the Southern Hemisphere, even if the attack were limited to the Northern Hemisphere. Effects of a nuclear attack during winter months

would typically be half as great as those in summer.

The estimates of temperature effects are based on sophisticated new modelling techniques that include the interaction between absorption of radiation and movement of the atmosphere, smoke represented as originating from regional sources and the use of scavenging rates determined by precipitation predicted by the model.

The new models can also lead to a lofting of the smoke due to warming that leads to a changed atmospheric temperature profile; the tropopause in the Northern Hemisphere can be lowered to 5 km, and the smoke above the new tropopause would not be efficiently scavenged, again leading to longer residence times (and decreased precipitation). The effect is more pronounced in summer because there is more solar energy to loft the smoke.

SCOPE has made new estimates of the quantity of smoke that would probably be produced in mass fires resulting from a 6,000 MT exchange. It arrives at an estimate of from 30 to 150 million tonnes once early scavenging has removed between 30 and 50 per cent of gross input; about 30 million tonnes of the remaining total would be strongly light-absorbent amorphous elemental carbon which, if spread over the Northern Hemisphere, could reduce insolation at the ground by 90 per cent. But the SCOPE authors admit that the effects of early scavenging represent one of the greatest sources of uncertainty, and the parameters are acknowledged as simplistic. Because of the strong concentration of fossil fuels and related products around cities and towns, the near-total burnout of less than one hundred of the

world's largest industrialized areas would cause the burning of 25–30 per cent of the combustible materials of the developed world.

The SCOPE study explicitly disavows the term "nuclear winter", which it says is too restricting for the multitude of possible effects of a nuclear war. The study makes some assessment of the effect of toxic substances other than smoke likely to be produced in a nuclear exchange. Nitrogen oxides, which may destroy atmospheric ozone, are seen as significant. And there are new gamma-ray dose estimates that suggest that under the same 6,000 MT exchange (about half the world's nuclear arsenal), lethal external gamma-ray doses (assuming no protective action) would cover 7 per cent of the land area of the United States, Europe and the Soviet Union. If spent nuclear fuel rods were distributed in the environment, the dose could be trebled. SCOPE also considers the effects of electromagnetic pulses (EMPs) caused by nuclear explosions.

SCOPE breaks new ground in its assessment of the biological consequences of cooling effects. The report considered the effects for a range of possible attack scenarios. Stress due to lowered temperatures, lowered light intensities and increased levels of ultraviolet light is found to have very serious consequences for agricultural systems in particular and, when taken in conjunction with levels of food in store in different countries, threatens starvation on a large scale. Even rather modest decreases in average temperatures — say 5° C for a growing season — would, for example, make the cultivation of wheat in Canada impossible. Many staple food crops, especially rice, are extremely vulnerable to temperature shifts, particularly during the growing season. The report recommends urgent research to clarify the effects of stress on crops as well as a comprehensive programme on the physics of smoke plumes and scavenging.

Tim Beardsley

Soviet missing person

Washington

THE SCOPE report on the environmental consequences of nuclear war pays special tribute to the contribution of Vladimir Aleksandrov, the Soviet nuclear winter researcher who mysteriously disappeared in Madrid earlier this year. Asked about Aleksandrov at the press conference announcing the SCOPE report, Sir Frederick Warner of the University of Essex, chairman of the project steering committee, said that Aleksandrov's whereabouts remained a mystery and a source of concern. He immediately offered the question to N. Lukyanov, who (like Aleksandrov before his disappearance) works at the Computer Centre of the Academy of Sciences of the USSR in Moscow. Lukyanov

said he was an old friend of Aleksandrov's but could add nothing to Sir Frederick's statement other than "we miss him".

The absence of information about Aleksandrov did not of course prevent speculation in the corridors of the US National Academy of Sciences, which was host to the 1985 SCOPE general assembly. One theory was that Aleksandrov, who by one account had unexplained conversations with Swedish scientists immediately before his disappearance, had planned to defect to the West. According to this theory, his disappearance was a pre-emptive move by Soviet intelligence services; Aleksandrov's whereabouts are believed to be unknown to the intelligence services of any Western country.

Tim Beardsley

Nuclear war

Other ways to reduce risks

Washington

THERE is much that can be done to reduce the risks of nuclear war independently of high-level strategic negotiations between the superpowers, according to the American Association for the Advancement of Science committee on science, arms control and national security. At a congressional seminar held here last week, the committee suggested several possible courses of action, including developing better crisis control mechanisms, strengthening conventional military forces, "confidence building", avoiding provocative strategic policies and trying to reach agreement on the reduction of nuclear weapons.

Congress last year passed a non-binding resolution calling for the establishment of nuclear risk-reduction centres. The feasibility of the centres was investigated by a group led by Senators Sam Nunn (Democrat, Georgia) and John Warner (Republican, Virginia) and used information from studies at Harvard, Stanford and Georgetown Universities in making specific recommendations. Last month, the Reagan administration formally agreed with the proposals and the concept was outlined to Mr Gorbachev by a group of senators who visited Moscow earlier this month. The Soviet reaction is reported to be "positive" and it is possible that the proposals will be included in the formal negotiations in November.

Nuclear risk reduction centres would maintain a 24-hour watch on events that might lead to a nuclear conflict and would be linked directly to relevant political and military authorities. One function of the centres would be to provide fast information about accidental or unauthorized launch of nuclear missiles, others would be the exchange of information about safety devices, joint planning during terrorist incidents, cooperation in any nuclear detonation by a third party and exchange of information about potentially threatening or misleading military activities, such as manoeuvres. The centres would not have a role in the resolution of "genuine" crises or in the negotiation of agreements, but would provide an institutionalized mechanism for a dialogue on how the United States and the Soviet Union perceive nuclear weapons risks independently of political negotiations. It is suggested that the centres be built in the two capital cities and linked via sophisticated real-time communications; there would be a liaison officer from each embassy. The centres, if established, would begin "modestly" with a narrowly defined function and would gradually evolve more complex roles; for example, a continuously updated database on nuclear forces could be maintained. Exactly how the centres would function and their res-

possibilities await their appearance on the negotiation agenda.

The School of Government at Harvard University has produced an *Agenda for Avoiding Nuclear War* by lessening the chance of errors or accident. The proposals include the maintenance of a credible nuclear and conventional military deterrent, crisis management training for personnel involved in making decisions under pressure and establishing procedures for termination of crises. Many of these suggestions have been incorporated in a resolution passed recently by Congress, which requires a report on the proposals from the State Department by January 1986. A similar resolution requiring the

Department of Defense (DoD) to assess the suggestions is likely to be passed later this month.

Further specific recommendations made to congressmen include the establishment of an annual "risk of nuclear war" report, requiring DoD and other relevant agencies to assess the effects of new weapons systems not only on arms control but also on the broad array of factors that affect the risks of nuclear war. Others are more funding for intelligence collection and modernization of communications and sanctions against non-nuclear weapons states that try to acquire nuclear weapons without imposing adequate safeguards. Representatives of the committee on science, arms control and national security say they are pleased by the interest shown by congressmen in their proposals.

Maxine Clarke

Nuclear winter

Mechanics of SCOPE report

Washington

THE Scientific Committee on Problems of the Environment (SCOPE) pulled no punches in presenting the principal conclusions of its report on environmental consequences of nuclear war in Washington last week.

The report, not yet formally published, is the product of a study launched in June 1982. The work has been carried out by a series of workshops, involving more than 300 physical scientists and biologists, and organized by a steering committee directed by Sir Frederick Warner and based at the University of Essex.

Mark Harwell of Cornell University, a principal author of the section of the report dealing with biological effects, said the key conclusion was that the indirect effects of a plausible nuclear war, principally starvation due to crop failures, might kill up to 4,000 million people, besides the several hundred millions who might die from direct effects. He invoked mass starvation in Ethiopia and the Sudan as more appropriate images of the aftermath of nuclear war than Hiroshima and Nagasaki, and estimated that more might die in a non-combatant nation such as India than in the Soviet Union and the United States put together.

Although the SCOPE report does not address policy questions, the SCOPE steering committee makes no secret of its wish that policy lessons be drawn from it. Sir Frederick Warner said that "anybody who thinks they can read this and not draw policy conclusions is making a big mistake".

Through SCOPE's parent body, the International Council of Scientific Unions (ICSU), efforts are being made to ensure that the striking conclusions from the new study are conveyed to the highest levels of government. The Prime Minister of India, Rajiv Gandhi, is already known to have a

special interest in the subject, and the report has been directed as well to officials in the governments of Australia, the United States, the Soviet Union and Britain.

The US Secretary of Defense, Caspar Weinberger, has said little on the subject of "nuclear winter", the usual name for large-scale cooling effects following a nuclear war. But an internal State Department memorandum to Secretary of State George Schultz dated 16 August 1984 says that "the implications for US policy of the nuclear winter theory as it is being argued by Turco, Toon, Ackerman, Pollack and Sagan (in *Science* 222, 1283; 1983) could be profound if the administration-sponsored studies agree with Turco *et al.*'s conclusions and/or if, by default, congressional and public attitudes are moulded by those results".

The US National Academy of Sciences study on the subject, published at the end of 1984, supported the "clear possibility" that the climatic consequences of a nuclear war might include a severe cooling of the kind described.

For last week's launching, SCOPE took the unusual step of hiring a public relations organization, the Center for the Consequences of Nuclear War, to publicize the report to the Washington press. The report will be formally published later this year, and a final version is nearing completion with support from private US foundations. ICSU has established a group to look at the conclusions and decide what research should be done next.

Some of the report's authors urged the audience at a scientific presentation of the results last week to use the report to oppose the use of nuclear weapons. And a Japanese SCOPE delegate, Yasuo Shimazu of Nagoya University, circulated an impassioned plea to delegates urging them to take active steps to prevent nuclear war from starting.

Tim Beardsley

Ariane launcher

Run of successes is broken

THE fifteenth launch of Ariane, Europe's expendable space launcher, ended as a further expense for space insurers last week. In the presence of the French President, M. François Mitterrand, the rocket had to be blown up over French Guiana with two communications satellites on board. The bill, according to Lloyd's of London, which had arranged the insurance of the satellites, will amount to some \$170 million.

But "there's no need to panic" over the future of the European space effort, a spokesman for Arianespace, the company that builds Ariane, said the day after the accident. According to Klaus Iserland, deputy director-general of Arianespace for technical matters, "the opposite situation" has prevailed for a few months: panic in the United States over satellite losses by the space shuttle and fears that Europe might sweep the board. "This one failure may reduce their panic a little but it's not suddenly going to reverse things."

The exact reason for failure in the Ariane 3 rocket is still unknown, so the effect of the loss on the Ariane launch programme is unpredictable. All that is known is that something went wrong with the ignition of the third-stage motor. Real-time data from the motor indicated the beginnings of ignition — a pressure build-up in the combustion chamber — and ignition was announced, but a few seconds later it was clear that the motor had not in fact ignited. The payload and third stage were then calculated to be on target for "a populated area", according to the European Space Agency, whose communications satellite ECS3 was aboard, and ground crew at Kourou then pressed the destruct button.

Now, teams are at work in Kourou and at the headquarters of Arianespace at Evry, near Paris, to analyse every hundredth of a second of telemetry from the launch vehicle to determine exactly what went wrong (which may take a few days) and why (which could take weeks). Identical third stages have been used on four previous Ariane 3 launches without any apparent problems, said Iserland.

The effect on the Ariane launch programme depends on the fault. If pinned down to a single faulty component, perhaps one which turns out simply to be from a poor batch, "we could be back very quickly" says Iserland. If, on the other hand, the problem is more complicated and requires design modifications and ground testing of the motor, the return of Ariane could take months.

Meanwhile, insurers who were tearing their hair out over space losses even before the latest failure may now be more difficult to convince that Ariane is really more reliable than the space shuttle. But Arianespace is going to continue to argue

that insurers should not attempt to recover losses on shuttle insurance by high premiums on Ariane. Lloyd's plans to wait to see the reaction when the next "risk" [a satellite to insure] comes on the market.

Robert Walgate

• The European Space Agency (ESA) claimed last week that it could meet its commitments to provide the European telecommunications satellite organiza-

tion, Eutelsat, with four communications satellites, despite the loss of ECS3 on Ariane. A fourth satellite, ECS4, is under development for launch in early 1987, and ESA says the programme could be accelerated to get ECS4 in space by mid-1986. A fifth satellite in the ECS series, a spare, will now also be made ready for production as ECS5. The programme has already been voted and funded "so ESA can meet its obligations without any difficulty." Launch costs for ECS5 will come from the \$65 million insurance that ESA took out on ECS3. □

Berne Convention

Britain lags on protection

BATS are well protected in Britain. They have it so good, in fact, that if other animals were equally well protected under the Wildlife and Countryside Act, Britain would have an excellent track record of enforcing the Berne Convention, which is the principal international agreement monitoring Europe's endangered wildlife. Instead, according to a report published last week by Wildlife Link, a consortium of environmentalists, Britain is failing to provide adequate habitat or protection for the animals it has promised to conserve.

"Britain voluntarily signed the Convention in 1979", says World Wildlife Fund consultant Simon Lyster. "They knew what they were letting themselves in for." The European Wildlife Convention, signed in Berne, outlines protection priorities for strictly protected plants (Appendix I), strictly protected animals (Appendix II) and protected animals (Appendix III). Britain's Wildlife and Countryside Act is intended to implement this agreement.

Corncrakes and merlins, great crested newts and smooth snakes: the Berne Convention is notable for the nonglamorous animals it protects. Providing that protection, however, requires complete surveys of animal habitats and population levels. France and West Germany, says Mr Keith Corbett of the British Herpetological Society, are running large computerized surveys of habitats, but the Nature Conservancy Council (NCC), which is entrusted with Appendix II protection, relies on unsystematic questionnaires. Habitat of very rare species, such as the corncrake, is not too difficult to identify; but "NCC is complaisant about great crested newts" because of their comparative numbers even though Britain has a large percentage of the European population.

The Wildlife Link report also points to the need to identify and defuse threats to wildlife populations protected by the convention. Otter numbers, for example, have not withstood pesticides, riverbank clearance and eel traps. At least 90 otters have been killed in eel traps in the past nine years, even though otter-excluding devices are easily installed. Identifying

threats, in other words, is only useful if the government offers incentives to farmers, companies and river authorities to modify their behaviour where necessary to protect endangered species.

Under the Wildlife and Countryside Act, the previously existing system of Sites of Special Scientific Interest (SSSI) is now being used and expanded as a means of upholding the Berne Convention. NCC must "renotify" an owner of an SSSI and identify the potential threats to the wildlife population there before the owner is obliged to negotiate with NCC over planned changes to the site. Only half the owners of Britain's 4,000 SSSIs have been renotified in the past three years, and it is unlikely that new SSSIs will be set up before this process is finished.

A principal problem of the SSSI system, according to the Wildlife Link report, is that SSSIs are chosen to protect representative types of environment — say, wetland or moor — and not species-specific habitats. What good, say environmentalists, is there in having the breeding areas of the merlin within SSSIs if its hunting grounds are not also protected? It would have been just as easy, claims Mr Lyster, to place the conifer plantations at issue in a different part of the uplands.

The sale of Appendix III species in Britain, such as frogs and smooth snakes, is controlled by a licensing system under the Department of the Environment. The system is "ludicrous", says Mr Corbett: the department has no idea how many animals are taken from the wild or where they come from. Three people in Britain applied for licences in the first half of 1984; yet well over 20,000 frogs were sold from Cornwall alone in 1983-84.

Britain's bats were lucky to have a successful lobby when the Wildlife and Countryside Act was passed, but any deficiencies of that act may not be the real problem. Britain has a good record with other international conventions on endangered species, but these affected such groups as jaguar-skin traders, not interests spread over the country. "The law is good enough", says Mr Lyster, "if the government wants it to be." **Elizabeth Collins**

ASAT

US weapons test criticized

Washington

THE United States' successful testing of its controversial new anti-satellite (ASAT) weapon against a satellite target last week took place despite the attempts of four Democratic congressmen and the Union of Concerned Scientists (UCS) to prevent the test.

The congressmen, together with UCS, had asked a US District Court judge to block the test because, they alleged, President Reagan has not yet met a congressional requirement that he certify the United States is negotiating in good faith with the Soviet Union to reach an agreement on limiting anti-satellite weapons. The case was dismissed by the judge because it was a political matter that should not be determined by the court.

The reaction from the Soviet Union to the test was predictably hostile: an embassy spokesman said the test would "inevitably entail negative consequences" for arms negotiations, especially since the Soviet Union has "exercised restraint" over ASAT weapons. The Soviet Union declared a moratorium on testing its own anti-satellite system two years ago but has threatened that a US decision to proceed with a satellite target test could lead to its own tests being re-started.

In its submission to the court last week, UCS included affidavits from two congressional observers at the US-Soviet arms control negotiations in Geneva. The observers testified that the United States has not responded to a Soviet proposal to ban testing of ASAT weapons, and that the United States has made no proposal of its own on the subject.

Critics of the test last week contend that it was conducted primarily for its political impact at Geneva, rather than for a genuine security need. The Soviet ASAT system is thought to be largely ineffective, and the target destroyed in last week's test was an out-of-service scientific satellite that was not properly instrumented to relay details of the weapon's performance. Technical problems continue to hold up a dedicated target satellite that was originally planned for the test.

Another factor that may explain the administration's enthusiasm to press ahead with the test, according to UCS, is that the technology of the ASAT's miniature homing vehicle is relevant to the Strategic Defense Initiative (star wars). The miniature homing vehicle, which destroys its target by the force of its impact, contains sophisticated infrared sensors that track the target and computers which fire 56 small steering rockets (each one no more than once). One component of the star wars research programme calls for "space based kinetic kill vehicles" that would be launched by rockets from an orbiting battle station at missile re-entry

vehicles; the sensor and steering technology involved would be essentially identical to that in the ASAT weapon.

UCS says, however, that last week's successful test against a single satellite should not be seen as evidence that a completely effective star wars missile defence is feasible. Most of the technical objec-

Supercomputers

UK machines oversubscribed

"THE Americans are seeing shadows that we do not" is the attitude of British university researchers and computer centres to US pressure for restricting Eastern bloc access to supercomputers. More to the point, the issue is irrelevant: the small numbers of supercomputers available for research as well as the costs of bidding for time on them create a self-imposed restriction that renders US pressure unnecessary.

Britain has two regional centres, in London and Manchester, for university research involving supercomputers. Funding for these centres was originally justified by the cost-effectiveness of doing certain types of research on machines working at speeds in the gigaflop range. In the short time since the London Cray computer came on-line in 1983 and the Manchester Cyber in 1984, the centres have become oversubscribed. Even if a new supercomputer centre is set up at the Rutherford Appleton Laboratory (see *Nature* 316, 569; 1985), the original centres could quickly fill the additional capacity both are now seeking.

Researchers wanting time on one of the supercomputers are screened through their universities or Britain's research councils, which then place bids for supercomputer access. Any government-imposed restrictions, therefore, would have to be aimed at all of Britain's universities, not just the two supercomputer centres, even though those centres make the final decisions about usage. Those lucky enough to be registered then use terminals in their own universities connected to the supercomputer (no instances are known of the password being broken). Time on the supercomputer, in other words, is so highly coveted that completing unauthorized work, particularly of any volume, on either machine is highly unlikely. The London and Manchester supercomputers are reserved for the use of the UK academic community and the work done on them, says Dr John Martin of the Manchester centre, has "no potential strategic implications whatsoever".

One possible avenue of US pressure is the inability of US companies, for whatever reason, to deliver computers on time or attempted imposition of restrictions on

tions to such a system concern the difficulties of tracking large numbers of re-entry vehicles and of distinguishing real re-entry vehicles from decoys.

The ASAT weapon tested last week probably has an altitude range of no more than 700 kilometres, as UCS also points out, which would put only 16 Soviet satellites within range of the weapon.

UCS is considering an appeal against last week's dismissal of its case.

Tim Beardsley

the machines' use. No restrictions or problems, says Dr Martin, were encountered during negotiations for the supercomputers already in place. "American companies are very keen to sell in Britain", says Dr Herbert of the Computer Board for Universities and Research Councils, and there is no evidence of intended pressure by US companies over machines for the proposed Rutherford centre. Nor is there any need for concern that supercomputer hardware will be resold outside Britain: such transfer, says Dr Herbert, "will never happen with supercomputers".

A concern of more relevance to university computer users, says Dr Martin, is the increased strength of restrictions placed by CoCom, the body coordinating Western high-technology export controls, on software produced either by universities or commercially. The blanket coverage of these restrictions is perceived as resulting in strict export limitations on software with no strategic implications at all.

Elizabeth Collins

More than super

According to recent reports in the Soviet media, the Soviet Union has developed a new "supercomputer", capable of performing 200 million operations a second (more than 10 times faster, it is claimed, than the US Cyber-73). The supercomputer, it is claimed, was developed entirely with Soviet resources, using a "basically new and more effective" data processing system, developed after the United States administration introduced restrictions on the sale of computers to the Soviet Union.

According to Moscow Radio (30 August), Western reports that the Soviet Union is lagging behind in electronics and computers are simply disinformation, circulated by those who seek to "discredit socialism and belittle its economic and scientific potential". Since electronics has become the key to industrial development, the commentator said, it must be clear to everyone that the Soviet Union, which accounts for one-fifth of the world's industrial production, cannot let itself be dependent on electronics imports.

Vera Rich

Nuclear technology transfer

Troubles for US-China pact?

Washington

THE transfer of nuclear technologies from the United States to China would provide short-term benefits but could pose risks to the United States by the end of the century if relations between the two countries were to deteriorate, according to a technical memorandum issued last week by the Office of Technology Assessment (OTA)*. The report was part of the evidence presented at a congressional hearing on the US nuclear energy cooperation agreement with China, which establishes a basis for cooperation between the United States and China in peaceful applications of nuclear energy; it is not a commitment to supply technology and expertise.

Increased trade, improved relations between the two countries and increased world energy supplies are the immediate benefits of technology transfer, OTA says. Exports to China could be worth several thousand million dollars in the next decade. OTA thinks it unlikely that China will find US reactor technology useful in nuclear weapons programmes, but expresses concern about China re-exporting to other countries and the disposition of spent fuel. Declarations by China's leaders that they will not assist others to develop nuclear weapons lack the force of a written treaty. But Ambassador Richard T. Kennedy, US negotiator of the agreement, said last week that China has made "significant new statements of its non-proliferation policy" and that he is satisfied that these policies are "consistent with our own basic views".

OTA warns, however, that China's nuclear submarines could benefit from "exposure to sophisticated nuclear industrial practices" in making possible quieter, more reliable and more powerful vessels. But the effects of a pact would be limited, says OTA, as China already has the reprocessing and enrichment technology that could create proliferation dangers and is importing from other countries the expertise that could be useful in improving weapons. Most energy technologies have no direct military applications, and the US-China pact expressly excludes the transfer of enrichment and reprocessing.

OTA was asked to produce the report by the Senate Banking Committee and a subcommittee of the House Energy and Commerce Committee. Senator Jake Garn (Republican, Utah), chairman of the former, strongly criticized the report for concentrating on the benefits to China and ignoring questions such as the future impact of technology transfer on US security and the effects of a strong China on countries in the region allied to the United States. Representative Ed Markey (Democrat, Massachusetts), a member of the House of Representatives' committee, questions whether the agreement will be

of significant value for US trade because of competition with French, German, British and Japanese companies already bidding for existing contracts, and because China is likely to become self-reliant in nuclear technologies very quickly. Markey criticized the vagueness of the agreement over the exchange of information and US visits to nuclear facilities in China.

Markey is attempting to show that the agreement violates the Nuclear Non-Proliferation Act of 1978 but it is unlikely that he will be able to stop Congress ratify-

ing the pact. If he proves he is right, President Reagan must resubmit the agreement to Congress with a waiver for those parts of it that do not meet statutory requirements. Congress would then have to approve of the pact by a joint resolution, which would give it special status. Other countries with technology transfer agreements with the United States would not then be able to use the US-China pact as a precedent for renegotiation. More probably, Markey's bid will fail and the pact will be ratified by default 90 days after its introduction by the President to Congress, February next. **Maxine Clarke**

*Energy Technology Transfer to China, US Government Printing Office, Washington, DC 20402; \$4.50.

Poland

Government denies pollution

CLAIMS that Poland is "the most ecologically threatened country in Europe" were denied last week by the Polish government press spokesman, Mr Jerzy Urban. The disputed report, he said, which was produced by the Polish Academy of Sciences' Committee on the Chemical Sciences is "not official", and its "alarmist and exaggerated" formulations are merely the views of "individual contributors".

Urban's criticism was ostensibly aimed, not at the report itself, but at an article based on it, which appeared in the daily *Zyccie Warszawy* (Warsaw Life) on 30 August. Refuting the article during his regular Tuesday press conference with foreign journalists, Urban claimed that Poland had "threatened areas", but no "areas of ecological disaster". Moreover, he said, the report did not reflect the government's commitment to the protection of the environment. The data used, he alleged, derived from the 1970s and even earlier.

In fact, the data of the academy's report cover the period up to the end of 1983. In a bold and, for present-day Poland, unusual move, the editors of *Zyccie Warszawy* next day published a leading article saying that although the government spokesman had access to more data than did its journalists (an allusion to the tacit reimposition of censorship in matters relating to the environment), a report presented to the Sejm (Parliament) in June by the head of the government's Office for Environmental Protection had "presented the current situation without beating around the bush and without concealing the serious dangers prevalent in certain areas".

The dismissal of the academy report as merely the views of certain individuals inevitably raises doubts as to the Polish government's willingness to listen to experts whose message fails to accord with the government viewpoint. In fact the academy document follows closely that of the report to the Sejm, even to the definition of 23 ecological "disaster areas". Other points raised by the academy include:

- A "shocking" increase in the number of retarded children in Upper Silesia, due to the increased concentration of heavy metals, particularly lead, in the environment.
- Air pollution in Poland is now the worst in Europe; of 1,066 factories whose emissions have been classified as environmentally harmful, only half have fixed maximum emission levels while 304 of the worst emitters have no set limits at all.
- The volume of automobile exhaust fumes is disproportionately high compared to the number of vehicles and, in spite of the severe fuel shortage, Polish vehicles burn twice as much fuel per kilometre as the world average.
- Poland has the highest proportion of smokers in the world, and the cigarette quality is particularly low, so that tobacco is a major source of indoor pollution.
- Of the 4,700 million m³ per year of waste water requiring treatment, almost half is discharged into rivers, lakes and the sea without treatment and another 30 per cent receives only mechanical treatment. More than 50 per cent of Polish cities, including Warsaw, have no sewage works. Seepage from urban and industrial effluent and the excessive use of fertilizers has led to massive degradation of ground waters.

The academy's report also claims that food crops are heavily polluted by industrial emissions. Close to urban industrial complexes, the concentration of cadmium is as much as 220 times the natural level; near the Boleskaw and Miasteczko Slaskie metallurgical works, the concentration of lead in market garden produce reached 230 mg per kilo (the maximum permissible dose is 3 mg per person per week).

According to *Zyccie Warszawy*, the team who prepared the report did so "to represent the strategic problems and suggest necessary action... with a measure of faith in its usefulness, despite the fact that examples from the recent past do not warrant such optimism". Mr Urban's unequivocal rejection of their report suggests, alas, that such "optimism" is still unjustified. **Vera Rich**

NERC scheme criticized

SIR—Sir Ronald Fisher once remarked that research students are the most important people in a university, so that the mechanism used by research councils for the award of research studentships is bound to be of great interest. The Natural Environment Research Council (NERC) has recently published a report of a committee under Professor R.J. Berry to review its own much criticized “project method” (Report of the Research Studentship Allocation Review Committee, November 1984).

By contrast, the Science and Engineering Research Council (SERC) has always allocated (as NERC formerly did) a number of studentships to individual university departments. Undergraduates apply to departments, are chosen for their potential quality and themselves choose the field of research in which they would like to work and thus their supervisor. NERC's scheme, introduced some years ago, gives a central role to its four training awards committees, each of which receives from each department in its field a list of 10–20 projects. The geological sciences training awards committee, for example, receives about 600 projects to which, in a two-day meeting, it allocates about 100 studentships. Two practical difficulties have emerged. Students may wish to work in a department on a project that happens not that year to have been designated, while the department may not be able to find anyone that year who wishes to work on the projects chosen.

There are more fundamental objections to the scheme. It removes from university departments to these committees an essential educational role. I have also argued (*Nature* 271, 704; 1978 and *The Observatory* 102, 164; 1982) that the ratio of geophysical to geological studentships is too small both for scientific advance and for national needs. The reason appears to be that there are many more geology than geophysics departments in universities and geophysics is still mistakenly equated with subdivisions of geology, such as stratigraphy, rather than being considered as an equal discipline: the Training Awards Committee consists of about 16 members, only one of whom is essentially a geophysicist. So the more fundamental aspects of geophysics as compared with those aspects closely related to crustal geology are under-represented in the projects chosen.

The Berry Committee consisted of the present and former chairmen of the geological science training awards committee, two former secretaries of the Medical and Agricultural Research Councils and Professor J.F. Dewey, who recently resigned from NERC's council (*Nature* 313, 730; 1985). It invited views on the project method and included a number of these in its report. From geology and geophysics

departments, there are 13 against and 4 for the scheme, but from the biological departments there are 10 for and 5 against. This asymmetry, not commented upon, may well arise because, for geoscience departments, the NERC scheme is the only way of supporting research students, whereas all biological departments also obtain research students from SERC.

The committee does not meet the criticisms of the NERC scheme raised by its respondents. For example, a geology professor states: “To anyone with a knowledge of the strength of geological sciences departments within the UK, the distribution of earmarked awards between them does not look sensible; major departments with high quality staff and vigorous research schools that are known to provide first-class research training are being underused for this purpose”. The committee does not meet this criticism, but instead discusses what constitutes research training.

Further, NERC's statement that “the primary objective of a research studentship is to enable the student to receive a good training in research, not to undertake a piece of novel research in a highly professional manner” was described by one geologist as badly worded, ambiguous and thoroughly misleading. The committee attempts to clarify this by specifying ten elements of research training, of which I quote three *verbatim*: “The definition of questions in order to be able to (a) disprove scientific hypotheses and (b) aid practical decisions; selection of materials for analysis, bearing in mind the possibilities of bias and the requirement for quality; the opportunity of obtaining an advanced understanding of an area of science via an in depth investigation of an appropriate research problem”.

Further obscure and ungrammatical verbiage occurs when the committee gets exercised on how to prevent PhD students untidily overrunning the allocated three years. It offers the following gems of wisdom: “These [procedures] include the drawing up of a realistic timetable which must be adhered to; the fact that three years is not a very long time; the importance of avoiding a slow start; the delays caused by failure to bring to a conclusion and distraction from the main line of enquiry. We commend these practices.”

Apparently some ill-natured people have raised the possibility that committee members stand a better chance than others of getting their projects approved. The tone of the report, so far unshakably complacent, becomes somewhat agitated on this point. “The evidence of inequality was not conclusive. . . the statistics are contradictory and open to different interpretations. . . Council takes a very serious view. . . an important area which should be carefully monitored in the future.”

These mysterious statistics are not produced, but no matter; “members (of the awarding committee) are drawn from amongst the best research workers and a statistically higher than average award rate is to be expected”!

Finally, the committee uses the venerable tactic of putting up a lot of Aunt Sallies, “numerous misconceptions” as it calls them. “The student is unimportant; the system is complex. . . time wasting. . . difficult to understand. . . disregards best students. . . disregards students who are innovative, favours particular supervisors and/or departments — committee membership is self perpetuating; topics are very narrow.” I cannot imagine any member of a university staff making such crude and sweeping criticisms except in moments of exasperation (which I am reliably told sometimes occur in dealings with NERC). The Berry committee has of course no difficulty in knocking them down, and comes to the conclusion that, except for minor alterations, no fundamental change in the scheme should be made. Dr J.C. Bowman, secretary of NERC, expresses in a preface “great pleasure in commending this report to all concerned with postgraduate training and research and sciences of the natural environment”, finding it “gratifying”. In NERC's eyes the scheme's virtue, as the report reveals, is that it can be represented to Whitehall as showing that university research is being guided in “relevant” directions.

Many will think this is not a recommendation and this “effortful, cumbersome, expensive and unsatisfactory scheme” should be abandoned and the considerable savings in cost used to support additional students.

KEITH RUNCORN

School of Physics,
The University,
Newcastle upon Tyne NE1 7RU, UK

Naked beauties

SIR—We read with naive interest but great expectation the article by D.J. Miller, “Naked beauty seen at last”, under the heading of “Particle physics”, which purports to explain why a beauty quark, otherwise known as a bottom, is difficult to perceive in its naked state (*sic*) (*Nature* 316, 681; 1985). We wonder whether other biologists reading this article similarly found it tantalizing, hilariously funny and totally incomprehensible. Do particle physicists have the same response to anti-idiotype networks (compare Mitchison's article in the same issue of *Nature*, p. 676)?

KATHY WEILBAECHER
MELVYN F. GREAVES

Leukaemia Research Fund Centre
at the Institute of Cancer Research,
Chester Beatty Laboratories,
Fulham Road,
London SW3 6JB, UK

DNA as a kind of solid

Molecules with helical pretensions such as DNA may be the result of a phase transition, in the strict sense, from a more disordered state. But when will the process be calculable?

THE textbooks say that double-stranded DNA molecules make a double helix, and that large globular protein molecules contain well-ordered helical structures (pieces of alpha helix) which are held together by less well ordered structures. No doubt the textbooks are right, at least for much of the time. But it is also well known that DNA molecules cannot function as templates for transcription while they are tightly bound as helices. For some of the time, the DNA molecules must be partly unwound. So much is generally accepted (and taught), but on the understanding that the helix is the normal resting state.

There is naturally much more that might be said. The complementary strands of DNA molecules are held together specifically by well-identified hydrogen bonds; by contrast, the forces that hold together the successive turns in a helically twisted molecule are, to say the best of them, loosely defined. People tend to build molecular models that allow the closest packing compatible with the molecular geometry and arrive at structures which turn out to be not very different from those suggested by X-ray diffraction. The double-strandedness is energetically understandable, the helical packing less so for DNA (but it is a rather different tale with the stretches of alpha helix in globular protein molecules).

So why do double-stranded DNA molecules finish up as helices? At bottom, the answer cannot be very different from that which explains why molecules of water vapour will condense at zero degrees centigrade or thereabouts, into the molecular constituents of solid ice. The trade-off is that between the decreased potential energy of a molecule trapped by attractive forces and what might be called the heat of disorder, the absolute temperature multiplied by the entropy.

For many familiar kinds of phase transitions, that from a vapour to a solid for example, the basis for understanding and even calculating the transition is familiar. Calculate the free energy of the two phases as a function of the temperature, ideally in terms of the properties of the constituent atoms and molecules, whereupon the crossover (if any) will mark the point of the phase transition. Now, and not before time, there is an argument bearing on the question whether the formation of helical DNA is a phase transformation in this sense.

The good news is that there is a formal calculation of the transformation of a disordered polymer molecule into a helix in which, as it happens, none of the parameters has been calculated from first principles. The other side of the coin is that the calculation, due to Nigel Goldenfeld of the University of California at Santa Barbara and J.W. Halley of the University of Minneapolis (*Phys. Rev. Lett.* **55**, 730; 1985), while providing an ingenious framework within which to describe systems of polymer molecules which are helical in parts and differently constituted elsewhere, also shows in passing how far there is to go.

The objective is to calculate the thermodynamics of a system of molecules each of which can in principle exist in two quite different configurations. Individual molecules will then be found in which stretches of one configuration are separated from each other by stretches of the other. In general, the lengths of these alternating stretches will be random, and all possible permutations may exist. The thermodynamic properties of the system are accessible if it is possible to calculate the partition function Z , which is nothing but the sum of terms such as $\exp(-E/KT)$, where K is Boltzmann's constant, T the temperature and E the energy of some state of the molecule as a whole. But the sum must be expanded over all possible states of the molecule, and this is usually where the snags arise. Generalities such as these are of course familiar from the textbooks.

Goldenfeld and Halley have made the problem tractable by introducing the assumption that the energy of a particular molecular arrangement, with alternating random lengths of the two configurations, is simply the sum of the energies of the consecutive segments. Among other things, this entails the assumption that energy states of the molecule as a whole are of negligible importance. In particular, low-frequency vibrations involving a whole polymer are assumed not to matter. With these simplifications, the problem

then boils down to calculating the energy states of random lengths of polymer in the alternative configurations (and of being able to stomach the algebra that follows).

The outcome is nevertheless encouraging. Goldenfeld and Halley suppose that the alternative configurations are respectively helical and straight. Each chemical unit of polymer molecule in the second form is supposed to involve a certain potential energy penalty, perhaps made up for by the extra degrees of freedom that it enjoys. Apart from the potential energy penalty entailed, the energy states of the helical and straight configurations differ only in the vibrational energy states that are accessible to them.

For the sake of argument, the energy states of the helical sections are approximated by those of a rigid rod. The outcome is that the energy of the rod-like sections is determined linearly by a combination of the length and of its logarithm, but that the free energy of the remaining sections is simply proportional to length. The free energy curves will plainly cross each other if the coefficients, which depend on the temperature, have suitable values, so that there must be a set of circumstances in which the transition to a helical structure is a straightforward phase transition, with a latent heat and so on.

The calculation has the great virtue of showing clearly that phase transitions in polymers are just as real as some experiments suggest. But it also illustrates what needs to be done to fill out the model. At least for the polymers occurring in biological systems, water relations are evidently crucial, partly because they will affect the molecular configurations by hydrogen bonding, and partly because the vibrations of water molecules will modify the vibrational spectra of both the coiled and uncoiled versions as well as serving as a means of transferring heat from one place in the molecule to another. But getting even that far quickly will be a far from simple achievement for whoever is successful.

John Maddox

Biological manuscripts

Contributors are reminded that, with the transfer of the Biological Sciences Editor of *Nature* to the Washington office, they should in future send **four copies of all manuscripts** to be considered for publication either (as at present) to the London office or to Washington.

Neurology

Predicting Parkinson's disease

from Solomon H. Snyder and Robert J. D'Amato

MANY researchers have been pessimistic about ever finding specific molecular abnormalities accounting for such important neurological disturbances as Parkinson's disease. Recent studies of N-methyl-4-phenyltetrahydropyridine (MPTP), a simple phenylpyridine derivative that elicits parkinsonism¹, have stimulated speculation that specific environmental toxins may underline most cases of Parkinson's disease. On page 246 of this issue, Calne *et al.*² report evidence suggesting that exposure to MPTP results in depletion of dopamine neurones in the striatum, the characteristic lesion of Parkinson's disease, in symptomless subjects. This raises the possibility that exposure to such toxins may predispose people to develop Parkinson's disease later in life.

MPTP is a byproduct in the synthesis of a meperidine-related opiate marketed illegally as synthetic heroin. Numerous cases of severe parkinsonism have been reported among individuals who have injected themselves with this drug^{3,4}. Monkeys receiving MPTP develop typical parkinsonian symptoms and have selective destruction of the nigrostriatal dopamine pathway⁵. MPTP binds with high affinity to receptor-like sites⁶, which appear to be identical to the enzyme monoamine oxidase B, found on glial cells in a number of discrete brain regions⁷⁻¹⁰. Monoamine oxidase transforms MPTP, whose pyridine ring is only partially unsaturated, into N-methyl-4-phenylpyridine (MPP⁺), which has a fully unsaturated pyridine ring and quaternary nitrogen (see figure). This transformation is required to make MPTP neurotoxic, since pretreatment of animals with inhibitors of monoamine oxidase prevents the induction of neurotoxicity by MPTP in monkeys¹¹ and mice¹². Why is MPP⁺ neurotoxic, while MPTP, which is closely similar in structure, is not? Unlike MPTP, MPP⁺ is accumulated via the pump-like catecholamine uptake system into dopamine neurones, which can concentrate it several thousand times above the levels in the surrounding extracellular space¹³.

Conceivably, environmental toxins similar to MPP⁺ or MPTP could cause naturally occurring (idiopathic) Parkinson's disease. One need not be exposed to the high levels of the agent obtained by intravenous injection: a thirty-seven-year-old chemist developed Parkinson's disease after working in a laboratory with MPTP¹⁴. Possibly such toxins elicit only a limited destruction of dopamine neurones and the natural aging process accounts for the rest¹⁵. In normal subjects dopamine levels decline with age by about 13 per cent per decade¹⁶; clinical symptoms of

parkinsonism begin to appear when depletion reaches about 70 per cent. Exposure to toxins could accelerate this process. Thus the four hundred or more individuals known to have injected synthetic heroin contaminated with MPTP but who are presently asymptomatic may be at risk of developing parkinsonism as they age.

Calne *et al.*² have now used positron emission tomography (PET scanning) to determine whether there is destruction of dopamine neurones in four such asymptomatic individuals. They compared these scans with those from six patients with idiopathic Parkinson's disease and seven normal subjects. PET scanning, using fluorine-18 labelled 6-fluoro-L-dopa that is taken up by nerve terminals and converted into 6-fluorodopamine, enables direct images to be made of the dopamine innervation of the corpus striatum. Calne *et al.* find a statistically significant decline in the formation and retention of ¹⁸F-dopamine in the four MPTP subjects, although not as great as in patients with severe idiopathic Parkinson's disease.

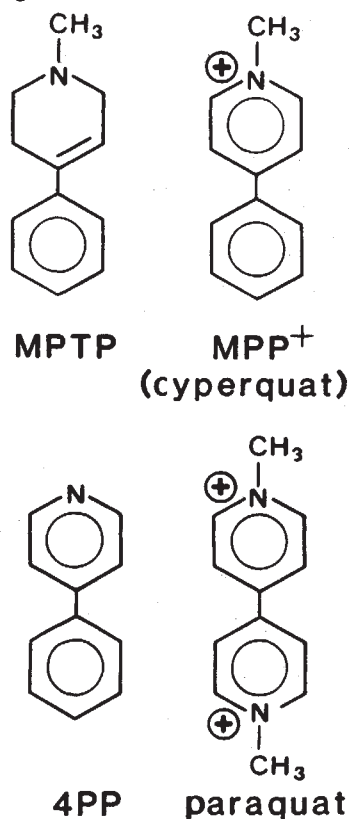
Because only four subjects were evaluated, one must be cautious about drawing strong conclusions. Moreover, two of the

four were receiving the opiate methadone and the other two had received narcotics until three months before the investigation. Opiates exert a variety of influences on dopamine metabolism that might have influenced the present results¹⁷. Despite these caveats, Calne *et al.* have probably detected a genuine loss of dopamine terminals in asymptomatic individuals exposed several years earlier to MPTP. As PET scanning and similar techniques become more widely available, predicting idiopathic Parkinson's disease in asymptomatic patients may become clinically feasible, although accompanied by ethical problems.

The present study provokes further thought about the role of environmental factors in the aetiology of idiopathic Parkinson's disease. If industrial pollutants are the culprit, one would anticipate variations in incidence in different countries and changes in incidence over the years but most epidemiological studies reveal a similar prevalence of parkinsonism in various countries and all decades of the twentieth century¹⁸. Many of these studies have focused on Northern Europe, but recent investigations show different incidences in the Mediterranean countries¹⁹, and in Quebec a high correlation between the disease incidence and pesticide use has been demonstrated²⁰. Lastly, some studies show a lower incidence of Parkinson's disease in black people²¹ and others report variations in the incidence of a type of parkinsonism in villages in the South Pacific separated only by a few miles²².

What environmental toxins are likely candidates? Hundreds of pyridines, similar in general structure to MPTP, exist in the environment, especially in foods. 4-phenylpyridine, which occurs in pepper-mint, spearmint, tea and industrial sources^{23,24}, produces a similar depletion of dopamine-related biochemical markers in PC12 cell cultures and in intact mice²⁵. MPP⁺ has been commercially marketed as a herbicide under the name cyperquat and the well-known herbicide paraquat structurally resembles MPTP (see figure).

One particularly tantalizing epidemiological finding bears on the role of pyridines in Parkinson's disease. Several well-designed studies show a reduced incidence of Parkinson's disease in smokers^{26,27}; contrary studies have generally employed fewer subjects²⁸. The inverse correlation between smoking and Parkinson's disease cannot, however, be attributed to a decrease in smoking after symptoms appear. Cigarette smoke contains approximately 100 distinct pyridines, each present at concentrations of several parts per million²⁹. Conceivably, these pyridines could compete with an MPP⁺-like toxin for uptake into dopamine neurones, thus blocking the action of the toxin, just as drugs that block the dopamine-uptake process prevent MPTP and MPP⁺ from damaging dopamine neurones¹³. Alternatively, carbon monoxide in cigarette smoke may



The molecular structures of N-methyl-4-phenyltetrahydropyridine (MPTP), its metabolite N-methyl-4-phenylpyridine (MPP⁺), the environmentally occurring analogue 4-phenylpyridine and the herbicide paraquat.

detoxify free radicals formed from environmental toxins¹⁴.

Pyridines are only one class of chemicals that could be toxic to dopamine neurones. The key requirement is that they or one of their metabolites, like MPP⁺, should be selectively concentrated in dopamine neurones. The present search for an environmental molecule involved in the aetiology of Parkinson's disease may or may not be successful but it is to be hoped that the quest will reawaken an interest in seeking specific molecular abnormalities associated with other major neurological and psychiatric disorders. □

1. Snyder, S.H. *Nature News & Views* **311**, 514 (1984).
2. Calne, D.B. *et al.* *Nature* **317**, 246 (1985).
3. Davis, G.C. *et al.* *Psychiatry Res.* **1**, 249 (1979).
4. Langston, J.W., Ballard, P., Tetrud, J.W. & Irwin, I. *Science* **219**, 979 (1983).
5. Burns, R.S. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **80**, 4546 (1983).
6. Javitch, J.A., Uhl, G.R. & Snyder, S.H. *Proc. natn. Acad. Sci. U.S.A.* **81**, 4591 (1984).
7. Wiecek, C.M., Parsons, B. & Rainbow, T.C. *Eur. J. Pharmacol.* **98**, 453 (1984).
8. Parsons, B. & Rainbow, T.C. *Eur. J. Pharmacol.* **102**, 375 (1984).
9. Levitt, P., Pintar, J.E. & Breakefield, X.O. *Proc. natn. Acad. Sci. U.S.A.* **79**, 6385 (1982).
10. Pintar, J.E. *et al.* *Brain Res.* **276**, 127 (1983).
11. Langston, J.W., Irwin, I. & Langston, E.B. *Science* **225**, 1480 (1984).

12. Heikkilä, R.E., Manzino, L., Cabbat, F.S. & Duvoisin, R.C. *Nature* **311**, 467 (1984).
13. Javitch, J.A., D'Amato, R.J., Strittmatter, S.M. & Snyder, S.H. *Proc. natn. Acad. Sci. U.S.A.* **82**, 2173 (1985).
14. Langston, J.W. & Ballard, P.A. *New Engl. med. J.* **309**, 310 (1983).
15. Calne, D.B. & Langston, J.W. *Lancet* (ii) 1457 (1983).
16. Reiderer, P. & Wuketich, S. *J. Neural Transm.* **39**, 257 (1976).
17. Blasig, J. in *Developments in Opiate Research* (ed. Herz, A.) (Marcel Dekker, New York, 1978).
18. Marttila, R. & Rinne, U. *J. Neural Transm.* **51**, 135 (1981).
19. Rosati, G. *et al.* *Neurology* **30**, 250 (1980).
20. Barbeau, A. *8th Int. Symp. Parkinson's Disease* (New York, June 9–12, 1985).
21. Kessler, I. *Am. J. Epid.* **95**, 308; **96**, 242 (1972).
22. Yanagihara, R., Garruto, R. & Gajdusek, C.D. *Ann. Neurol.* **13**, 79 (1983).
23. Sakurai, K., Takahashi, K. & Yoshida, T. *Agric. Biol. Chem.* **47**, 2307 (1983).
24. Maga, J. *J. Agric. Food Chem.* **29**, 896 (1974).
25. Snyder, S.H., D'Amato, R., Nye, J. & Javitch, J.A. in *Proc. MPTP Symp.* (eds Markey, S. P., Castagnoli, N., Kopin, I. J. & Trevor, A.) (Academic, in the press).
26. Godwin, R.B., Lee, P.N., Marmot, M.G. & Stern, G.M. *J. Neurol. neurosurg. Psych.* **45**, 577 (1982).
27. Baumann, R., Jameson, H., McKean, H., Houck, D. & Weisberg, L. *Neurology* **30**, 839 (1980).
28. Rajput, A. *Can. J. Neurol. Sci.* **11**, 156 (1984).
29. Heckman, R. & Best, F. *Tobacco Sci.* **25**, 33 (1981).

Solomon H. Snyder and Robert J. D'Amato are in the Departments of Neuroscience, Pharmacology and Experimental Therapeutics, and Psychiatry and Behavioral Sciences, Johns Hopkins University School of Medicine, 725 North Wolfe Street, Baltimore, Maryland 21205, USA.

Palaeontology

Dinosaurs that fill the gaps

from Michael J. Benton

DINOSAURS arose in the late Triassic, 225 Myr ago; some of the early forms have become quite familiar, such as small carnivorous theropods (for example, *Coelophysis*), large herbivorous prosauropods (for example, *Plateosaurus*) and small herbivorous ornithomimids (for example, *Fabrosaurus*). For a long time, it was thought that these early species were followed by a gap in the fossil record spanning about 30 Myr, after which there appear a selection of completely different dinosaurs. But the gap has now been filled by the re-dating of some dinosaur beds and because of new discoveries in the early middle Jurassic of Europe, India and China.

The first dinosaur faunas after the gap occur in the second half of the middle Jurassic in Europe (about 170 Myr BP). For example, the Bathonian horizons near Oxford, England have produced rare skeletons of the theropod *Megalosaurus*, as well as the giant herbivorous sauropod *Cetiosaurus* and the armoured stegosaur *Lexovisaurus*. Better known are the late Jurassic (about 155 Myr BP) dinosaurs *Allosaurus*, *Brontosaurus*, *Diplodocus* and *Stegosaurus* from North America.

The re-dating exercise has extended the age of dinosaur-bearing beds that were once placed at, or just below, the Triassic/Jurassic boundary, right through to the end of the early Jurassic (187 Myr BP). A combination of biostratigraphical evi-

dence (fossil pollen and spores, fish and footprints) and radiometrically determined ages of interbedded lavas has shown that these units cover a much wider time span than had been suspected^{1,2}. Intervals of time that were formerly considered to be devoid of dinosaurs are now known to have been populated over almost all the world by small theropods such as *Syntarsus*, medium-to-large prosauropods such as *Massospondylus* and *Euskelosaurus*, and small ornithomimids such as *Fabrosaurus* and *Heterodontosaurus*.

The dinosaur faunas of the early Jurassic have also been extended by new finds, particularly of early sauropods. *Barapasaurus* from India³ and *Vulcanodon* from Zimbabwe⁴ (see figure) have been interpreted as intermediate between the prosauropods and the sauropods. These were large animals, 8–10 metres long, that probably walked habitually on all fours. Another sauropod of similar age is *Ohmdenosaurus* from Germany⁵, represented by only a few limb-bones.

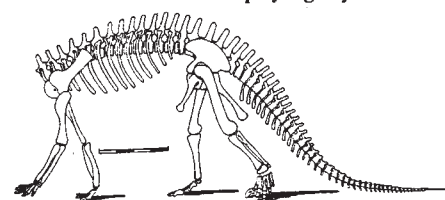
The middle Jurassic record of dinosaurs has been dramatically supplemented by new finds from China. In the past few years, several important new dinosaur specimens have been excavated from the Xiaoshaximiao Formation of the Sichuan Basin (age: Bathonian–Callovian, 176–163 Myr BP).

A skull and partial skeleton of one of

the earliest stegosaurs has been named *Huayangosaurus*⁶. The skull is primitive in several respects, and the armour consists of bony plates and spines. Until now, the earliest known stegosaurs came from Europe, but the new Chinese find indicates an eastern Asian origin for the group. Two sauropods have also been described from this formation: *Shunosaurus*^{7,8} and *Datousaurus*⁸. More than 10 skeletons of *Shunosaurus* have been found; this species was 7–8 m long, whereas *Datousaurus* was 14 m long. Both these sauropods show similarities to the English *Cetiosaurus*, which is of about the same age, but the skeletal remains of the Chinese animals are much more complete. These specimens, together with *Barapasaurus* and *Vulcanodon*, provide important new evidence that the sauropods arose from the prosauropods of the late Triassic and early Jurassic. Some palaeontologists had argued that the sauropods arose directly from Triassic thecodontians (the basal archosaur group), but that aspect of the theory that dinosaurs arose polyphyletically is now contradicted.

The fourth new dinosaur from the middle Jurassic of China is the ornithomimid *Xiaosaurus*⁹, a bipedal herbivore only one metre long. It might be related to *Fabrosaurus*, but the remains are fragmentary. In addition a carnivorous dinosaur *Xuanhanosaurus*¹⁰ has been identified on the basis of several limb-bones and vertebrae. The hand has the typical blade-like claws of a predator and, in general, this theropod seems to resemble the English *Megalosaurus*, which is of about the same age.

Palaeontologists, who have often been handicapped in their studies by gaps in the fossil record, will rejoice at these important new discoveries. Already, they are improving our understanding of the faunas and of dinosaur phylogeny. □



Reconstruction of *Vulcanodon karibaensis* Raath. Trunk and ilium adapted from *Barapasaurus*. Scale bar, 1 m.

1. Olsen, P. E., McCune, A. R. & Thomson, K. S. *Am. J. Sci.* **282**, 1 (1982).
2. Olsen, P. E. & Galton, P. M. *Palaeontol. afr.* **25**, 87 (1984).
3. Jain, S. L., Kuttu, T. S., Roy-Chowdhury, T. & Chatterjee, S. *Proc. R. Soc. B* **188**, 221 (1975); in *4th International Gondwana Symposium 1977* (eds Laskar, B. & Raja Rao, C. S.) 204 (Hindustan, Delhi, 1979).
4. Cooper, M. R. *Palaeontol. afr.* **25**, 203 (1984).
5. Wild, R. *Stung. Beitr. Naturk.* **B41**, 1 (1978).
6. Dong, Z., Tang, Z. & Zhou, S. *Vertebr. palasiat.* **20**, 83 (1982).
7. Dong, Z., Zhou, S. & Shang, Y. *Palaeont. sin.* **C23**, 1 (1983).
8. Dong, Z. & Tang, Z. *Vertebr. palasiat.* **22**, 69 (1984).
9. Dong, Z. & Tang, Z. *Vertebr. palasiat.* **21**, 168 (1983).
10. Dong, Z. *Vertebr. palasiat.* **22**, 213 (1984).

Michael J. Benton is in the Department of Geology, The Queen's University of Belfast, Belfast BT7 1NN, Northern Ireland, UK.

Astrophysics

Cepheids and cosmic distance

from David A. Hanes

OBSERVATIONAL cosmologists await with excitement the launch of the Hubble Space Telescope, which bids fair to revolutionize our view of the extragalactic Universe and, in so doing, eventually to resolve the vexed question of the correct cosmic distance scale. Nevertheless, despite optimistic predictions that the whole problem will be resolved in the course of a few days' observations (Maddox, *J. News and Views Nature* **313**, 347; 1985), answers seem likely to require years of observation and analysis. This is especially the case as much of the attention has centred on the projected study of cepheid variable stars, long an indispensable tool of observational cosmologists but so far undetectable in any but the nearest galaxies.

But now, B. Madore and W. Freedman (*Astron. J.* **90**, 1104; 1985) have taken a fresh look at the problem and have devised new efficient procedures which eliminate much of the tedium that has so far seemed unavoidable in the study of cepheids. Given the extremely high costs of operating the Space Telescope, the savings are of great practical importance, offering the promise of a swifter result that will contribute to the resolution of a long-standing controversy.

Cepheid variable stars are in some respects excellent distance indicators. They are intrinsically luminous and thus can be detected in remote galaxies; they draw attention to themselves by varying regularly in brightness with periods which range from a few days to a few hundred days; and they obey a well-defined period-luminosity law, with the intrinsically bright cepheids varying more slowly than their less luminous brethren. Their use is straightforward in principle. Repeated observations reveal the variability and allow the delineation of the light curve, whence comes a secure mean apparent luminosity and the period. The latter implies an absolute luminosity, once the period-luminosity law has been calibrated by nearby cepheids of known distance.

Erratum

Apologies are due to our readers and to Ira Herskowitz for an unfortunate typographical error in his article "Yeast as the universal cell" (22 August, page 678). The first sentence of the fourth paragraph should have read: "The yeast spindle is a tripartite structure, consisting of microtubules that join at one end of the centromere (the yeast version of the kinetochore of the chromosome) and at the other end to the microtubule organizing centre, the spindle pole body (analogous to the centrosome)". In the published article 'centrosome' appeared as 'centromere', thus exacerbating the terminological confusion the author was trying to clarify.

and the difference between apparent and absolute luminosity yields the distance.

Practice is more difficult. Historically, cepheids have most often been studied through the use of blue-sensitive photographic plates, which assists detection because cepheid amplitudes are largest at short wavelengths. Moreover, photographic plates can be made uniform on large scales, and are thus ideal for survey work in nearby galaxies (although they lack the sensitivity of modern electronic detectors). But each cepheid must be well studied throughout its cycle, and hundreds of long-exposure plates may be needed to delineate the light curves; thus some programmes have taken decades to complete (Sandage, A.R. *Astrophys. J.* **166**, 13; 1971). Principally for this reason, extragalactic cepheids have been studied in only a few nearby galaxies. Given the small field of view of the Space Telescope and other urgent demands on it, a repetition of such programmes is neither practicable nor desirable.

Madore and his co-workers (*IAU Colloq.* **82**; 1985) have already nicely simplified one aspect of the use of cepheids, by relying on the fact that, at red and near-infrared wavelength, cepheids vary with only small amplitude. Thus a single infrared 'snapshot' of a cepheid allows a precise distance estimate, provided of course that the period is already known. Moreover, the effects of interstellar obscuration and different stellar heavy element abundance are reduced at the longer wavelengths, so that certain systematic errors may be eliminated. Even so, in as yet unexplored galaxies, the need to carry out long-term surveys both to identify the cepheids and to determine their periods seems to be one of the prices which must be paid.

Or so it appeared until Madore and Freedman's most recent investigation. The question they posed was this: how can one determine the period of a cepheid without many observations at various phases? Their use of a simple analogy aids understanding their conclusion. Imagine an idealized case in which all cepheids have the same amplitude in astronomical magnitudes (relative luminosity). Suppose further that the light curve of a cepheid is a pure saw-tooth function; that is, each cepheid declines linearly in magnitude over its full period, then instantly brightens back to maximum luminosity. In this picture, long-period cepheids have slower decline rates than do short-period cepheids; moreover, a pair of brightness measurements only days apart should yield the period by a simple inversion.

Real cepheids, of course, are not so

obliging. Their amplitudes are not all the same. Moreover, while their light curves are generally not sinusoidal, neither are they perfectly saw-toothed; and it is the lower-amplitude cepheids that have the more nearly sinusoidal behaviour. One complication therefore is that a pair of randomly placed observations may straddle the turnover in the light curve or fall on the rising branch, so that the derived period is quite incorrect. However, Madore and Freedman show that, by happy coincidence, some of these effects cancel out or can be unambiguously taken into account; they predict that a large fraction of the detected cepheids will yield adequately precise periods from a single pair of brightness measurements.

There are some other complications. Although cepheids are likely to be in the majority, any star field will include variable stars of different kinds, for which the treatment is quite invalid and misleading. Moreover, any particular cepheid may have a binary companion diluting its light. But Madore and Freedman show that only slight extensions of the proposed methods — the use of several rather than two spaced observations at a few wavelengths — will resolve most of these difficulties.

Indeed Madore and Freedman describe encouraging simulated experiments in which they analysed pairs of observations separated by six or seven days, randomly drawn from the massive compilations of the photometry of the well-studied cepheid variables in the Small Magellanic Cloud, the nearest external galaxy outside the Milky Way. These observations yield reliable periods for the predicted fraction of the cepheids and produce a quite reasonable period-luminosity relationship. Although the sample is restricted, without interloping variables of other kinds, the results lend impressive support to the proposed techniques. Madore and Freedman next plan to study nearby galaxies with ground-based telescopes to quantify the accuracy attainable in actual observations.

Will these efficient techniques alone

100 years ago

THE NEW STAR IN ANDROMEDA

WE have received the following important communication from Lord Rosse:

SINCE my communication of September 8 our books have been searched for information on the past history of the nucleus of the Andromeda nebula. I subjoin in full the entries bearing upon the question whether the "new star" is now seen for the first time, or is a variable now shining out with abnormal brilliancy. The latter would appear to be the case. The nebula was frequently observed in past years with the 6-foot reflector and measures made. These measures being too few in number for a proper survey of the nebula, publication was postponed in 1878, and the details of configuration of the nebulosity have not appeared such as to merit a monograph. ROSSE

From *Nature* **32**, 465, 17 September 1885

permit a quick resolution of the contentious question of the correct cosmic distance scale? Sadly, the answer is no. A major part of the difference between the so-called 'long' and 'short' distance scales (de Vaucouleurs, *G. Observatory* **102**, 178; 1982) stems from disagreements in the calibration of the intrinsic period-luminosity relationship for cepheids. The calibration depends on those few cepheids of known distance in our own Galaxy and is sensitive to the assumed membership of variables in loose star clusters and to the

effects of interstellar and circumstellar obscuration. Madore and Freedman's techniques will shed no light on these fundamental questions of calibration. Yet they will extend our ability to derive precise relative distances for external galaxies in reasonable time periods and must be welcomed as a major contribution to the effective use of a new instrument. □

David A. Hanes is in the Department of Physics, Queen's University, Kingston K7L 3N6, Canada.

Materials science

A spongy way to new ceramics

from Paul Calvert

INDUSTRIAL ceramics such as alumina, zirconia or titania are normally prepared from a 'green body' of compacted powder that is sintered at 1,500°C or more, slowly converting by solid-state diffusion to a dense solid. But the increasing use of ceramics in microelectronics and their potential as engine parts have imposed requirements of uniformity and strength that are difficult to meet with this process while using crude powders as starting

the grains and pores are the same size before sintering. This egalitarian state should sinter rapidly to a dense, fine-grained, strong ceramic. The success of this approach has recently been demonstrated for silica by Sacks and Tseng¹. As an amorphous material silica is not such an interesting ceramic as alumina or zirconia, but the work goes a long way to vindicating the concept. This approach depends on the ability to produce uniform sub-micron spheres, which was described for silica by Stober and colleagues² and involves the slow hydrolysis of silicon tetroxide in alcoholic solution. Subsequently, this has been extended to hydrolysis of other metal alkoxides, including zirconium and titanium, but it has so far proved very difficult to produce single-size particles in the 0.1–1 µm range in quantities to make the process commercially attractive.

One joke in the ceramics field is that the only reason for working with spheres is that we are not clever enough to make cubes. It would appear that we might take some lessons from the sponges. Their skeletons are a particularly impressive example of the use of silica, which forms spicules of about one micron diameter in a

huge variety of shapes including umbrellas, commas and double-ended anchors. They are formed intracellularly within a vesicle surrounded by a membrane, the silicalemma³, and the shape is determined by deposition from dilute silicic acid on a protein filament of 0.2 µm diameter. Li and Volcani⁴ have recently reviewed the deposition of silica in diatoms, which appear to employ a similar mechanism: deposition takes place within the silicalemma, and polymeric (possibly polysaccharide) filaments act as templates for the formation of columns that fuse to form solid plates; four forms of silica can be distinguished.

Although the ceramics industry has no immediate requirement for umbrella-shaped particles, they might well find a use for them if they were available. The possibility of forming particles using a membrane to control the reaction and polymeric templates are ideas worthy of consideration. First the companies must convince themselves that the perfection and control available through such techniques will compensate for the much greater cost of powders.

An intriguing marriage of the natural and synthetic approaches currently being developed commercially is the production of silicon carbide fibres from rice hulls that contain up to 30 per cent by weight of silica. Pyrolysis in the absence of air results in the conversion of the silica to SiC fibres of about 0.3 µm diameter and 50 µm length that seem to have great potential as a reinforcement for ceramics and metals⁵. □

1. Sacks, M.D. & Tseng, T.-Y. *J. am. Ceram. Soc.* **67**, 526 (1984).
2. Stober, W., Fink, A. & Bohn, E. *J. Colloid & Interface Sci.* **26**, 62 (1968).
3. Hartmann, W.D. in *Silica and Siliceous Structures in Biological Systems* (eds Simpson, T.L. & Volcani, B.E.) (Springer, Berlin, 1981).
4. Li, G.W. & Volcani, B.E. *Phil. Trans. R. Soc. B304*, 519 (1984).
5. Sharma, N.K., Williams, W.S. & Zangvil, A.J. *J. am. Ceram. Soc.* **67**, 715 (1984).

Paul Calvert is in the School of Chemistry and Molecular Sciences, University of Sussex, Brighton BN1 9QJ, UK.

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Top surface of 0.5 µ spheres of compacted titania (E.A. Barringer, Massachusetts Institute of Technology).

materials, and companies are thinking seriously about making powders by processes similar to those sponges and diatoms use to produce skeletal silica.

For maximum strength the sintered body should have a very fine grain size and either no pores or, at least, no large pores. During sintering a 'rich get richer' law operates: the big pores in the compacted powder initially grow at the expense of the small ones, before shrinking slowly by vacancy diffusion; meanwhile the small grains are swallowed up by the large ones. Thus the final structure is normally 95–99 per cent material with a few large pores and a large average grain size (10–100 µm) — the opposite of what is wanted.

A philosophy argued strongly by H.K. Bowen of the Massachusetts Institute of Technology is that these problems could be avoided by starting from an 'ideal' green state. The aim is to produce sub-micron spherical particles of uniform size and then pack them into a regular lattice or into dense random packing so that all

IMAGE
UNAVAILABLE
FOR
COPYRIGHT
REASONS

Scanning electron micrographs of fossil diatoms from upper Cretaceous deposits in western Siberia. The siliceous exoskeletons of these unicellular algae consist of two valves and a variable number of girdle bands. The species illustrated are joined together in chains; *a* shows two valves, one from each of two adjacent cells. In *b* the interlocking spines and the fine perforations through the silica wall are seen in more detail. *c* is a detail of a closely related genus with partial occlusions of the pores. From Ross, R. and Sims, P.A. *Bull. Br. Mus. nat. Hist. (Bot)* **13**, 277 (1985) by courtesy of the Trustees of the British Museum (Natural History).

Drosophila genetics

Patterns of differentiation

from P. W. Ingham

How do genes control pattern? Some answers to this question are now beginning to emerge as the cloning and characterization of *Drosophila* pattern-forming genes gathers momentum. A cluster of genes comprising the bithorax complex^{1,2} plays a key role in the differentiation of various body parts of the fly. These genes are thought to act by controlling the expression of other genes, being selectively expressed in specific lineage units or compartments³. Support for this model at the molecular level is presented in two recent papers^{4,5} which describe the distribution of products of the *Ultrabithorax* gene (*Ubx*) in *Drosophila* embryos.

Ubx is one of the three major genes of the bithorax complex⁶. It is itself a genetically complex locus, subsuming four separately mutable functions which map to two large transcription units (Fig. 1a)^{2,7}. The larger of these, termed the *Ubx* unit, spans some 73 kilobases and is the site of all *Ubx* mutations which have been mapped so far. Next to this is the smaller *bxd* unit, defined by mutations which abolish only the *pbx* and *bxd* sub-functions of the locus (for definition, see legend to Fig. 1).

Studies by Akam and Martinez-Arias⁴ have concentrated on expression of the *Ubx* unit. Three major RNA species (3.2, 4.3 and 4.7 kilobases in size) are transcribed from the unit during the embryogenesis of *Drosophila*. Using a probe common to the upstream (5') end of all three, they find, by *in situ* hybridization, that *Ubx* is expressed from an early stage in embryogenesis. This expression is restricted to the posterior portion of the embryo, specifically to seven metameric units in the early segmented embryo (Fig. 1c), a finding in fairly close agreement with the predictions of E. B. Lewis's model of the bithorax complex¹. According to this model, each gene of the complex has a primary role in the development of a specific body segment. For *Ubx*, the major function would be in the metathoracic segment, and the minor in the abdominal segments A1 to A7.

Recently Struhl⁸ and Hayes *et al.*⁹ have proposed a modification to this model. On the basis of various new observations of mutant phenotypes they suggest that the domain of function of each bithorax gene is out of register with the body segments. Thus, each gene would control the development not of a single segment, but rather of a parasegment¹⁰, a metameric unit that comprises the posterior lineage compartment of one segment and the anterior compartment of the next. If true, this underlines the importance of compartments as the basic building blocks of the animal. Therefore, it is of consider-

able interest to see if the accumulation of *Ubx* transcripts coincides with segmental or parasegmental units. At first sight, the pattern of *Ubx* expression in the early embryo (Fig. 1c) seems at odds with this 'out of register' hypothesis as the region expressing *Ubx* is neatly bounded by ectodermal grooves that are conventionally thought to correspond to segment boundaries. However, recent evidence¹⁰ suggests that these grooves do in fact correspond to the anterior/posterior boundary within each segment, and thus delineate parasegments. If this is so the parasegmental requirement for *Ubx* inferred from genetic analysis is precisely matched by the pattern of *Ubx* expression that is observed *in situ*.

Independent confirmation of such parasegmental expression has been obtained by White and Wilcox^{5,11} and by Beachy *et al.*⁷. Using antibodies raised against a synthetic partial *Ubx* protein (derived from the common 5' region of the *Ubx* unit), they analysed the distribution of *Ubx* proteins in later embryonic stages by indirect immunofluorescence. These studies reveal a region of high *Ubx* protein expression in the central nervous system, which, although equivalent in width to a single segmental neuromere, appears out of phase with the segmental arrangement

of the neuromeres (Fig. 1d). White and Wilcox⁵ have now extended this analysis to the epidermis; here too, the boundaries of *Ubx* protein accumulation lie between the segment borders. Interestingly, the *Ubx* protein is localized in the nucleus, as expected of a protein involved in the regulation of gene expression.

The region showing the highest expression of *Ubx* products, parasegment 6 (the posterior compartment of the metathorax plus the anterior compartment of A1), corresponds to the domain which requires the *bxd* function of the *Ubx* gene (Fig. 1e)^{1,6}. The relationship between *Ubx* and *bxd* is puzzling, especially as all *bxd* mutations are known to map to a distinct transcription unit. Lewis's model explains the effect of *Ubx* mutations on the *bxd* function in terms of a *cis* inactivation of the *bxd* gene by the former. But the new results suggest that the *bxd* function is mediated by expression of the *Ubx* product. Thus, a more attractive hypothesis seems to be that the *bxd* unit specifically regulates the expression of *Ubx* in parasegment 6 (refs 4,7,12).

By preparing probes specific for the different *Ubx* transcripts, Akam and Martinez-Arias⁴ have discovered a complex pattern of expression of the three major embryonic messenger RNAs. One striking feature of this pattern is the absence of any obvious correlation between the different transcripts and the various functions encoded by the *Ubx* locus. For instance, the preferential expression of *Ubx* associated with the *bxd* function in parasegment 6 is not attributable to any

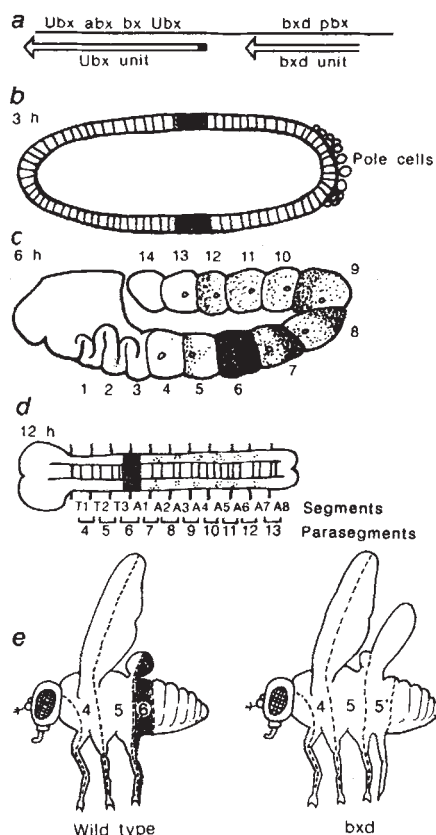


Fig. 1 a, the *Ubx* pseudo-allelic series consists of the *abx*, *bx*, *Ubx*, *bxd*, and *pbx* mutations which together define about 120 kilobases of DNA on chromosome 3R. This region gives rise to the *Ubx* and *bxd* transcripts. **b**, *Ubx* transcripts (shaded region) are first detected at the blastoderm stage of embryogenesis and are localized in a small region corresponding to the primordia of segments T3 and A1. **c**, about 3 h later a series of ectodermal grooves subdivide the embryo into metameric units; these correspond to parasegments — the posterior compartment of one segment and the anterior compartment of the next. The *Ubx* transcript is expressed at high levels in parasegment 6, the posterior compartment of T3 and anterior compartment of A1, and at lower levels in parasegments 7 and 12 and in part of 5. **d**, after a further 6 h, the segmental organization of the embryonic CNS is easily visualized — each segmental repeat is characterized by two lateral commissures and is separated by an intersegmental nerve. At this stage, high levels of *Ubx* protein can be detected in the posterior part of segment T3 and the anterior of A1 (parasegment 6); low levels of the protein are present in the posterior of segment T2 and in abdominal segments 2 to 7. **e**, this high level of *Ubx* expression in parasegment 6 seems to mediate the *bxd* function of the *Ubx* locus. In the adult, the parasegmental boundaries are defined by the anterior/posterior cell lineage boundaries in each segment (dotted lines). In *bxd* mutants, parasegment 6 becomes transformed into a duplication of parasegment 5.

particular species of transcript but rather reflects elevated levels of all three. In late embryonic stages, high levels of the 3.2- and 4.3-kilobase mRNAs are present in parasegment 6 of the CNS. In earlier stages, when the 4.3-kilobase message is barely detectable, parasegment 6 is preferentially labelled by probes specific for the 3.2- and 4.7-kilobase transcripts. Thus, no single *Ubx* transcript can be considered to mediate the *bxd* function.

Characterizing the disposition of *Ubx* products is only the first step towards elucidating the way in which the bithorax complex works. At present we can only speculate on the significance of the spatial and temporal patterns of expression of the different transcripts. One thing they do imply is the existence of a complex underlying regulatory mechanism. Some measure of this complexity is indicated by Jurgens¹³, who describes a class of genes which act synergistically in regulating the expression of bithorax genes. Using this synergistic interaction as an assay, he estimates that about 40 loci are required for the correct expression of bithorax complex genes, although the specificity of these requirements is not known.

One of the most intriguing aspects of bithorax regulation concerns the way in which the expression of genes such as *Ubx* is initiated in the appropriate regions of the embryo. Akam and Martinez-Arias show that *Ubx* transcripts are first detectable just before the cellular blastoderm stage of embryogenesis when membranes are beginning to isolate each nucleus. By the time that cells have formed, high levels of the 4.7-kilobase transcript are present in a single zone which, from its position, probably corresponds to the primordium of parasegment 6 (Fig. 1b). As preferential accumulation of all three transcripts in parasegment 6 is a motif which characterizes *Ubx* expression throughout

embryogenesis, the early accumulation of the 4.7-kilobase transcript is clearly a key event in establishing the subsequent pattern of *Ubx* expression.

But how is this initial 4.7-kilobase expression regulated? Unfortunately, the phenotypic criteria used^{1,12,14} to identify regulatory genes have not permitted discrimination between genes controlling the initiation of *Ubx* expression and those required for its expression later in embryogenesis. *In situ* analysis of transcription now allows a direct assessment of the stage at which a particular mutation affects *Ubx* expression. Various considerations suggest that maternally acting genes are significant in the initiation of the bithorax complex genes. However, the need to co-ordinate the genetic programming of cells with the physical subdivision of the embryo into segmental units suggests that there is also an interaction between *Ubx* and those zygotically acting genes which control segmentation¹⁵. Such an interaction has already been proposed in the case of the segmentation gene *fushi tarazu*¹⁶. The way is now open for investigation of these possibilities. □

1. Lewis, E. B. *Nature* **276**, 565 (1978).
2. Bender, W. *et al. Science* **221**, 23 (1983).
3. Garcia-Bellido, A. *Ciba Fdn. Symp.* **29**, 161 (1975).
4. Akam, M. & Martinez-Arias, A. *EMBO J.* **4**, 1689 (1985).
5. White, R. A. H. & Wilcox, M. *EMBO J.* (in the press).
6. Sanchez-Herrero E. *et al. Nature* **313**, 108 (1985).
7. Beachy, P. A. *et al. Nature* **313**, 545 (1985).
8. Struhl, G. *Nature* **308**, 454 (1984).
9. Hayes, P. M. *et al. Proc. natn. Acad. Sci. U.S.A.* **87**, 545 (1984).
10. Martinez-Arias, A. & Lawrence, P. A. *Nature* **313**, 639 (1985).
11. White, R. A. H. & Wilcox, M. *Cell* **39**, 163 (1984).
12. Ingham, P. W. *Cell* **37**, 815 (1984).
13. Jurgens, G. *Nature* **316**, 153 (1985).
14. Struhl, G. *Nature* **293**, 36 (1981).
15. Nusslein-Volhard, C. & Wieschaus, E. *Nature* **287**, 795 (1980).
16. Laughan, A. & Scott, M. P. *Nature* **310**, 25 (1984).

P. W. Ingham is in the Mill Hill Laboratories, Imperial Cancer Research Fund, Burtonhole Lane, London NW7 1AD, UK.

Particle physics

Real and imaginary monopoles

from Richard A. Carrigan Jr

ON PAGE 234 of this issue a group at Imperial College London¹ report on a large-scale search for the elusive magnetic monopoles using an inductive detector. Though none has been found, these experiments have illuminated some of the ways that an inductive detector can mimic a monopole signal. This reflects a persistent concern that the original monopole "detection" by Cabrera, which caused great excitement amongst particle physicists and astrophysicists, may also have arisen from such an effect.

The last decade has indeed been exciting for monopole fans^{2,8}. The observation by 't Hooft and Polyakov that many important group theories contain super-heavy monopoles was shortly followed by

the discovery and subsequent dismissal of a candidate monopole event in a balloon-borne emulsion stack. In 1979, John Preskill brought the problem into sharp focus with his famous paradox — on the one hand the physics demanded roughly one monopole per proton in the Universe, but on the other hand this number of monopoles would have closed the Universe on a time scale that would have left only Bishop Ussher happy. This conundrum was shortly followed by the concept of monopole catalysis, the theory that a magnetic monopole can induce proton decay, so that a passing monopole would leave a trail of expiring protons in its wake.

Cold water was thrown on all this de-

licious excitement by Alan Guth with his concept of inflation. Guth's theory inflates the cosmos till one monopole or so resides in the Universe — surely a discouraging possibility for a hunter. With inflation, Guth could explain Preskill's paradox as well as the very homogeneous character of the Universe.

Guth's inflation might have discouraged most of the searchers had it not been for the appearance of a solitary signal in a single detector operated by Blas Cabrera at Stanford. Cabrera has a well-deserved reputation for building extremely good magnetic field shields, spaces essentially free of magnetic fields, at the 10^{-11} T level. Cabrera attached a 5 cm diameter current loop to a superconducting quantum interferometer device inside his multi-layer shield and waited for six months. On St Valentine's Day 1982, the device picked up a current step — exactly the unique signal expected from a monopole passing through the loop and pulling magnetic flux with it.

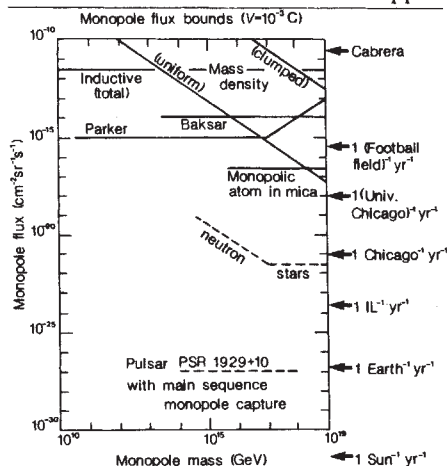
Strictly speaking Cabrera had no business looking for monopoles. Existing ionization detectors had already set much lower limits than he could reach for free cosmic monopoles based on an expected high-ionization signal. These experiments were open to a criticism: slowly moving massive monopoles would not be so strongly ionizing, so might have escaped the detector net. More telling was an astrophysical limit, the Parker bound: even a small taint of free monopoles in the Galaxy would have shorted out the galactic magnetic field, which would have happened were the monopole flux a factor of 10,000 – 100,000 times lower than the Cabrera event suggested. Nevertheless Cabrera's event was a very clean signal in a detector uniquely suited to a monopole search.

In short order, a three-front attack was mounted on the problem of monopole abundance. First, the underlying principles of the ionization detectors were re-examined and old and new detectors swung into action, quickly establishing plausible limits thousands of times lower than that suggested by Cabrera's candidate. Second, the astrophysicists were having a field day: assessing the Parker bound, looking at the effects of monopole catalysis on neutron stars, and generally setting astonishingly low numbers for monopole abundances. And finally, a third front was established with a new wave of inductive detectors including a larger Cabrera apparatus⁴, a detector at IBM¹⁰ and a Chicago-Michigan-Fermilab device¹¹.

The Imperial College result¹ is characteristic of this third front. The detector employs three conductor loops mounted inside a single superconducting shield, the largest with an area more than 50 times greater than Cabrera's original loop. This and other large detectors represent a compromise in design: multiple loops for

coincidences, but correspondingly less effective shielding and more inductive coupling between the shield and the loop.

But the problems with the shielding and inductive coupling may have been a blessing in disguise, for they have illuminated some of the ways that an inductive detector can mimic a monopole signal. Mechanical motion of the shield can move trapped



Current monopole limits drawing on Turner⁸ for astrophysical limits and Giacomelli⁶. Appropriate limits are taken at a monopole velocity of a thousandth of the speed of light. The inductive number is from Caplin *et al.*¹ and combines Imperial College, Chicago-Fermilab-Michigan¹¹, IBM¹⁰ and more recent Stanford upper limits. Baksar is the largest ionization detector. The lowest Earth-based limit assumes that a monopole will capture an atomic nucleus. Astrophysical catalysis limits are shown as dotted lines. (Catalysis cross-sections are assumed to be 10^{-28} cm² limit.) The Irvine - Michigan - Brookhaven multiple-catalysis limit is the same as the inductive limit. All experimental limits are 90 per cent confidence levels.

magnetic flux in and out of the detector¹. The bottom line seems to be that most step signals in the loops can be easily explained in terms of mechanical shocks as the helium level in the Dewar rises and falls. The original Cabrera event may have been generated in a similar way.

In any case, the Imperial College experiment, in concert with other negative results (including a new Cabrera detector run for 100 times the original area-time factor), gives an overall upper limit of 2×10^{-12} cm⁻² sr⁻¹ s⁻¹. This excludes the original Cabrera event but is still 1,000 times higher than the Parker bound.

So what is the future for inductive searches and other earth-based searches? The original impetus to the development of inductive detectors was stimulated by the Cabrera candidate and a suspicion that slow massive monopoles might not register in ionization detectors. The larger inductive detectors have illuminated some of the problems with these devices, while thorough analyses of the ionization properties of monopoles have shown that suitably efficient electronic detectors can be made. Indeed, many of the electronic monopole hunters are now combining forces to fill the Gran Sasso tunnel in Italy

with a detector that will reach and perhaps exceed the Parker bound⁶. This seems a worthwhile goal.

With Guth's inflation and the negative experimental results, are magnetic monopoles totally dead? Don't bet on it. There is as yet no completely satisfactory link between inflationary cosmology and particle physics. This is further complicated by the lack of experimental verification of the most straightforward grand unified theory. Newer physics — super-symmetry or superstrings — still can have built-in monopoles. An interesting exemplar¹² of a group with two sets of monopoles and two levels of symmetry breaking is SO(10). This might have non-catalysing monopoles with masses as low as 100 TeV. In the last year or so there has been intense discussion of TeV physics beyond the W.

Space medicine

The skeleton in space

from A. W. Goode and P. C. Rambaut

WITH astronauts now making multiple short trips and with the prospect of their undertaking recurrent 3-month tours of duty on the Space Station, an understanding of the influence of gravity on the skeleton becomes urgent. Inability to find appropriate preventative measures may lead to increased urinary calcium, calcium deposits in the kidney, renal stones and hypertension¹. Between middle and old age, approximately 15 per cent of skeletal tissue is lost², whereas post-menopausal women may lose up to 1 per cent per year³. These losses, added to the unresorted or irreversible space-flight losses, could lead to an increased risk of fractures for astronauts with advancing age. New procedures being developed may help to determine whether the changes in bone density observed in space are reversible.

The effects of space flight are frequently compared with the effects of paralysis or long-term bed-rest. In such situations the calcium loss seems to be self-limiting; however, remobilization in adult humans does not seem to correct completely the mass defect of the bones⁴. Whether a similar risk confronts the astronauts remains to be determined.

Undesirable changes in bone density in space flight were first confirmed in the cosmonauts of Vostoks 2 and 3 who were found to have increased levels of urinary calcium excretion⁵. During the following two decades of space flight more data from increasingly sophisticated testing methods have continued to document the loss of calcium from bones.

The two-man crews of the Gemini IV, V and VII missions sustained losses in bone density in the heel bone ranging from 2.9 to 15.1 per cent as measured by post-flight radiographic densitometry⁶. However, astronauts and cosmonauts of the Gemini

Perhaps one of these theories will even have a laboratory-sized monopole in it. □

1. Caplin, A.D., Guy, G.N., Hardiman, M., Park, J.G. & Schouten, J.C. *Nature* **317**, 234 (1985).
2. Carrigan, R.A. Jr & Trower, W.P. *Nature* **305**, 673 (1983).
3. Craigie, N.S., Goddard, P. & Hahm, W. (eds) *Monopoles in Quantum Field Theory* (World Scientific, Singapore, 1981).
4. Carrigan, R.A. Jr & Trower, W.P. (eds) *Magnetic Monopoles* (Plenum, New York, 1983).
5. Stone, J. (ed.) *Monopole 83* (Plenum, New York, 1984).
6. Giacomelli, G. *La Rivista del Nuovo Cim.* **7**, 1 (1984).
7. Preskill, J.A. *Rev. Nucl. Sci.* **34**, 461 (1984).
8. Turner, M.S. *Fermilab-Conf. 85/72A*, Proc. First Aspen Winter Conf. (in the press).
9. Gardner, R., Cabrera, B., Taber, M. & Huber, M. *Proc. 17th Int. Conf. Low Temp. Phys.*, 945 (North Holland, Amsterdam, 1984).
10. Tesche, C.D., Chi, C.C. Tsuei, & Chaudhari, *App. Phys. Lett.* **43**, 384 (1983).
11. Incandela, J. *et al. Phys. Rev. Lett.* **53**, 2067 (1984).
12. Weinberg, E.J., London, D. & Rosner, J.L. *Nucl. Phys. B* **236**, 90 (1984).

Richard A. Carrigan Jr is at the Fermi National Accelerator Laboratory, Batavia, Illinois 60510, USA.

VII and Soyuz 9 flights had no change in serum calcium levels but exhibited increased urinary calcium excretion. Post-flight data showed a decrease in calcaneal bone density of 8–10 per cent^{7,8}.

On the flights of Apollos 7 and 8, bone density changes measured by improved X-ray densitometry techniques⁹ showed some bone loss in the heel bone in five of the six astronauts¹⁰. ¹²⁵I-photon absorption measurements on astronauts in flights of Apollos 14, 15 and 16 (up to 12 days' duration) showed bone density changes to be 6.6 per cent in one astronaut and 7.3 per cent in another¹⁰. Furthermore, on the most recent lunar flight (Apollo 17), the estimated total body calcium loss was 0.2 per cent. For the first time, an apparent disturbance in the gastrointestinal absorption of calcium was noted with both urinary and fecal calcium excretion increasing during weightlessness¹¹.

Data from the 3 Skylab flights (of up to 84 days' duration) revealed that the quantity of calcium lost each day in the urine began to increase as soon as the crew members became weightless. During the first month, the loss was only 50 mg per day but it was over 200 mg per day by the end of the second month¹². In both Apollo and Skylab programmes the decrease in bone density occurred in the calcaneus, but not in the radius or ulna, confirming that mineral loss occurs primarily from weight-bearing bones in crew members on all space flights.

Increased bone turnover or resorption was seen in Skylab during the first 30–40 days of weightlessness and then tended to stabilize¹³. Fecal calcium output initially declined in-flight, then rose above pre-flight values. Following Skylab flights, calcium balance tended to become less negative, but did not reach equilibrium during

the 18-day post-flight observation period. The mineral density in Skylab astronauts 5 years after their flight is lower than before their flight¹⁴. Data from the Russian flights of the Salyut programme, which has included missions of up to 237 days in duration, reveal similar reductions in mineral content¹⁵. It has been suggested¹⁶ that these smaller than anticipated bone losses can be attributed to various countermeasures. A 'penguin' suit, worn almost continuously by cosmonauts in long-term missions¹⁷, is designed to provide a load of about 50 per cent of the body weight on the cosmonaut when in the normal earthbound posture. And maintenance of nutrient intakes at high levels and extensive exercise on the ergometer and the treadmill are also thought to reduce bone losses.

Losses of calcium of the magnitude observed in space are smaller than those seen in long-term immobility in patients. However, the study of skeletal change in space in the absence of disease offers a unique opportunity to understand mechanisms of skeletal integrity for the benefit of both astronaut and patient. □

1. Goode, A. W. *Nature* **283**, 525 (1980).
2. Trotter, M. I., Broman, G. E. & Peterson, R. R. *J. Bone Joint Surg.* **42-A**, 50 (1980).
3. Cohn, S. H., Vaswani, A., Zanzi, I., Ellis, K. J. *Am. J. Physiol.* **230**, 143 (1976).
4. Parfitt, A. M. *Acta Astronaut.* **82**, 1083 (1981).
5. David, H. *Missiles and Rockets*, 34 (1963).
6. Mack, P. B., LaChance, P. A., Vose, V. P. & Vogt, F. B. *Am. J. Roentg.* **100**, 503 (1967).
7. Lutwak, L., Whedon, G. D., LaChance, P. A., Reid, J. M. & Lipscomb, H. S. *J. clin. Endocr. Metab.* **29**, 1140 (1969).
8. Birykov, N. & Krasnykh, I. C. *Kosmicheskaya Biologiya i Meditsina* **4**, 42 (1970).
9. Mack, P. B. *Aerospace Med.* **42**, 828 (1971).
10. Rambaut, P. C., Smith, M. C., Mack, P. B. & Vogel, J. M. in *Results of Apollo NASA SP-368* (eds Johnston, R. S., Detlein, L. F. & Berry, C. A.) 303 (NASA, Washington DC, 1975).
11. Rambaut, P. C., Leach, C. S. & Johnson, P. C. *Metabolism* **182**, 62 (1975).
12. Rambaut, P. C. & Johnson, R. S. *Acta Astronaut.* **6**, 1313 (1979).
13. Leach, C. S., Rambaut, P. C. & DiFerranti, N. *Acta Astronaut.* **6**, 1323 (1979).
14. Tilton, F. E., Degioanni, J. C. & Schneider, V. S. *Aviat. Space Envir. Med.* **51**, 1209 (1980).
15. Stupakov, G. P., Kazaykin, V. S., Kozlovsky, A. O. & Korolev, V. V. *Cosmic Biol. Aerospace Med.* **18**, 33 (1984).
16. Gagenko, O. G., Genin, A. M. & Egorov, A. D. *Acta Astronaut.* **2**, 97 (1981).
17. Umansky, S. P. *Cosmonaut Equipment, Moscow* **5**, 59 (1982).

A. W. Goode is Assistant Director of the Surgical Unit, The London Hospital, London E1 1BB, UK; P. C. Rambaut is at NASA Headquarters, Washington, DC 20546, USA.

Gene modification

Targeting in mammalian cells

from Tom Maniatis

ALTHOUGH the feasibility of stably introducing cellular genes into mammalian cells was demonstrated several years ago¹, the targeting of genes to their normal chromosomal location has not been achieved until now. But on page 230 of this issue Smithies *et al.* report a series of elegant experiments that clearly demonstrate the feasibility of introducing a cloned human β -globin gene into its normal chromosomal location by homologous recombination in cultured cells². These experiments lay the groundwork for developing practical methods for the targeted modification of mammalian genes and they may also improve the prospects of human gene therapy.

When cloned DNA is introduced into cells by calcium phosphate precipitation³ or microinjection⁴, it is incorporated into the host genome at random sites by a mechanism that does not involve homologous recombination (reviewed in ref. 5). Although cloned DNA sequences recombine with each other when introduced into mammalian cells⁶⁻¹³ evidence for recombination between the introduced gene and its natural chromosomal counterpart has been difficult to obtain. The possibility that transfected histocompatibility genes can recombine with their endogenous counterparts has been suggested by two studies^{14,15}, but the identification and characterization of the products of such recombination events have not been reported. Other studies have demonstrated

recombination between a mutant selectable gene stably introduced into mammalian cells (an 'artificial target gene') and a homologous transfected gene carrying a mutation at another site within the gene^{10-13,16}, but direct evidence that these recombination events result in the site-specific integration of the cloned DNA was not obtained.

Smithies *et al.*² owe their success in targeting the human β -globin gene to its normal chromosomal location to a novel approach that incorporates several important features for optimizing the frequency of generation and the sensitivity of detection of targeted recombinant molecules. First, the donor or test plasmid contains DNA sequences that allow it to be discriminated both before and after integration from the normal chromosomal copy of the β -globin gene. Second, a cell line in which the β -globin gene is active is used as the recipient cell, because insertion of the plasmid into an inactive locus could turn off the expression of the linked selectable marker (the neomycin resistance gene, *neo*, which confers resistance to the antibiotic G418), and thus eliminate any true recombinant from the population of transformed cells. Third, the donor plasmid is cleaved within the β -globin 5'-flanking sequences by *Bst*XI to optimize the frequency of recombination, on the basis that cleavage of transforming DNA within the region of homology significantly increases the frequency of homologous recomb-

ination in yeast¹⁷ and mammalian cells¹⁸. Fourth, a bacteriophage lambda gene rescue procedure^{19,20} is used to identify and isolate the recombinant DNA fragments from a mixture of total genomic DNA through its association with a bacterial suppressor tRNA gene (*supF*).

As shown in Fig. 1, recombination between the donor plasmid and the chromosomal copy of the globin gene will generate a unique *Xba*I fragment of 7.7 kilobases that contains the *supF* gene and the intact β -globin gene plus 5'- and 3'-flanking sequences. This fragment can be rescued from total genomic DNA of the host cell by digesting with *Xba*I, ligating to an *Xba*I phage vector bearing two amber mutations, packaging the phage *in vitro*, and then plating on *Su*⁻ bacteria. Those recombinant phage that grow on *Su*⁻ bacteria will have acquired the *supF* gene and contain either randomly integrated globin genes or the targeted gene. Phage containing the targeted gene can be distinguished from the random inserts by the presence of the second intron (IVS 2) of the β -globin gene. In a typical experiment², approximately 10⁸ viable phage were recovered from the packaged DNA, 1,000 of these formed plaques on *Su*⁻ bacteria, and just one of these phage contained the targeted gene. The initial transformation frequency was 10⁻⁴ or 10⁻⁵, giving an overall homologous recombination frequency of approximately 10⁻⁷–10⁻⁸ of the cells exposed to the transforming DNA.

This low frequency of recombination immediately raised concerns such as whether the rare recombinant phage is the result of recombination during *in vitro* packaging of the phage DNA, or of contamination from laboratory phage stocks.

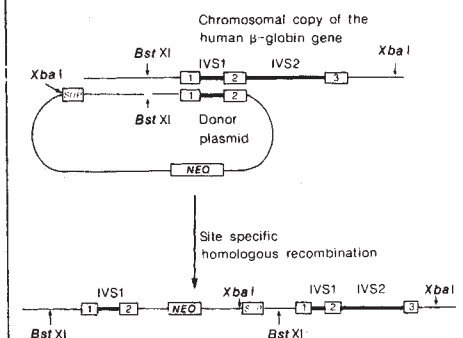


Fig. 1 Strategy for introducing cloned human β -globin gene into chromosomal DNA by homologous recombination. Numbered boxes are exons; thick lines are introns (IVS). Restriction endonuclease cleavage sites for the enzymes *Bst*XI and *Xba*I are indicated. Dashed lines indicate the site of recombination between chromosomal DNA and the donor plasmid.

To address these concerns, Smithies *et al.* found and investigated the cell line containing the targeted β -globin gene. Genomic-blotting analysis of this line revealed that the integrated donor plasmid had the structure shown in Fig. 1, proving that a homologous recombination event had indeed occurred.

Implications

Although the frequency of homologous recombination provided by this technique is presently too low for most applications, at least recombination can be observed and an assay is available to detect it. From this beginning it may now be possible to develop methods for increasing the frequency of homologous recombination.

One ultimate application of such techniques is to alter genes in their normal chromosomal context, which could be invaluable in the analysis of the mechanisms of gene expression in higher eukaryotes. As yet, it has not been possible to establish rules for the appropriate regulation of cloned genes introduced into mammalian cells. But numerous studies that show qualitative or quantitative differences between the expression of transfected genes and their endogenous counterparts suggest that the chromosomal context of some genes play an important role in their expression. For example, β -globin genes introduced into mouse erythroleukemia cells (MEL) at random chromosomal positions are appropriately regulated when the cells are induced to differentiate into erythroid cells, but the amount of β -globin mRNA produced is 10–100-fold less than from the endogenous globin genes, on a per gene basis^{21,22}. A more extreme example is the expression of the cloned human α -globin gene in cultured cells. This gene is regulated correctly when introduced into MEL cells by chromosome transfer methods²³ but is not regulated when it is introduced into the same cells by DNA transformation²⁴. Similarly, a number of genes are inappropriately regulated when introduced into the germ line of mice (see ref. 25 for review). On the other hand, there are examples in which cloned genes introduced into the germ lines of flies^{26–28} or mice²⁵ are regulated in a manner that is indistinguishable from that of the endogenous genes, irrespective of their location in the genome of the transformant. Thus, at least in certain cases, normal chromosomal location is not a prerequisite for normal gene expression.

Regardless of the reasons for the observed variability of expression of cloned genes in mammalian cells, it is clear that precise gene replacement would be preferable to random integration if it could be accomplished. In this regard, it is important to note that Smithies *et al.*² have achieved only the first step of precise gene replacement. As shown in Fig. 1, the integrated β -globin gene contains a partially duplicated β -globin gene, vector sequen-

ces and the selectable gene. Thus, the chromosomal context of the integrated β -globin gene differs significantly from the normal situation. This is of concern, since it is known that β -globin genes are inappropriately regulated when they are introduced into transgenic mice linked to vector sequences; removal of these sequences results in the normal expression of the gene even though it is integrated at random chromosomal locations^{29–31}. It would therefore be interesting to know whether the integrated β -globin gene in the cell line characterized by Smithies *et al.*² is appropriately regulated.

In contrast to mammalian cells, precise gene replacement has been a routine procedure in yeast for several years^{32–34}. As shown in Fig. 2, the first step in this process is analogous to that used by Smithies *et al.*². Homologous recombination between a wild-type gene in the donor plasmid (which contains *Ura3*⁺ as a selectable marker gene) and a chromosomal gene containing a mutation, results in the in-

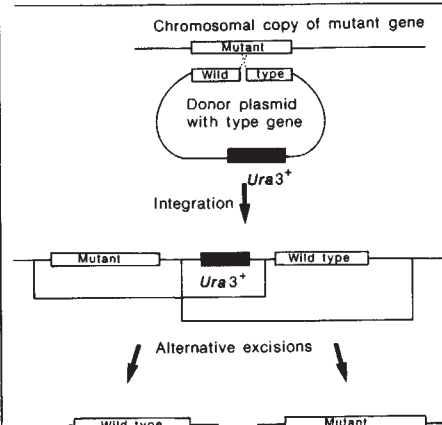


Fig. 2 Strategy used for precise gene replacement in yeast.

tegration of the wild-type copy of the gene adjacent to the mutant copy. A subsequent recombination event will lead to the excision of either the mutant or wild-type copy of the gene, depending on the site of crossing over. Now that the first step in gene replacement has been demonstrated in mammalian cells, it may be possible to achieve the second step by using a strategy similar to that used in yeast. However, the enormous background of random integration events in mammalian cells compared with that in yeast represents a major obstacle to the identification of mammalian cells in which recombination has occurred. Moreover, except in the occasional hemizygous cell line, only one of the two copies of the gene of interest will be altered in the diploid mammalian cell, presenting the difficulty of studying transcription of the altered gene in the presence of an unaltered gene. This problem also complicates gene disruption experiments of the type used so effectively in yeast³⁴. Thus, although it should now be possible to replace genes in mammalian cells, the technology in its present state does not

provide a practical method for studying gene expression.

To what extent does the successful gene targeting method devised by Smithies *et al.*² improve the prospects of gene therapy? The authors indicate that the ability to target genes may ultimately allow the modification of defective β -globin genes in the bone marrow stem cells from patients with β -thalassaemia of sickle-cell anaemia. But, as they note, enormous difficulties remain to be solved. For example, the frequency of transformation and homologous recombination is orders of magnitude too low to target genes in haematopoietic stem cells, which are present in very small numbers in bone marrow. In addition, it is not presently possible to select for homologous recombinants over the high background of random insertion events. Therefore, virtually all of the cloned β -globin genes introduced into stem cells would not be targeted.

Whether gene targeting methods or retrovirus vectors^{35,36} are eventually used for gene therapy, and whether gene therapy proves to be practical or beneficial, remains to be established (see ref. 37 for discussion). In any case, the accomplishments of Smithies *et al.*² represent an important first step in the *in situ* manipulation of mammalian genes. □

- Wigler, M. *et al.* *Cell* **16**, 777 (1979).
- Smithies, O. *et al.* *Nature* **317**, 230 (1985).
- Robins, D.M. *et al.* *Cell* **23**, 29 (1981).
- Capecci, M.R. *Cell* **22**, 479 (1980).
- Pellicer, A. *et al.* *Science* **209**, 1414 (1980).
- Folger *et al.* *Molec. cell Biol.* **2**, 1372 (1982).
- Liskay, R.M. & Stachelek, J.L. *Cell* **35**, 157 (1983).
- Small, J. & Scangos, G. *Science* **219**, 174 (1983).
- Subramani, S. & Berg, P. *Molec. cell Biol.* **3**, 1040 (1983).
- Folger, K., Thomas, K. & Capecci, M.R. *Cold Spring Harb. Symp. quant. Biol.* **49**, 123 (1984).
- Smithies, O. *et al.* *Cold Spring Harb. Symp. quant. Biol.* **49**, 161 (1984).
- Lin, F.L., Sperle, K. & Sternberg, N. *Cold Spring Harb. Symp. quant. Biol.* **49**, 139 (1984).
- Smith, A.J.H. & Berg, P. *Cold Spring Harb. Symp. quant. Biol.* **49**, 171 (1984).
- Goodenow, R. *et al.* *Nature* **301**, 388 (1983).
- Yoshie, O. *Proc. natn. Acad. Sci. U.S.A.* **81**, 649 (1984).
- Lin, F.L., Sperle, K. & Sternberg, N. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1391 (1985).
- Orr-Weaver, T.L., Szostak, J.W. & Rothstein, R.J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6354 (1981).
- Kucherlapati, R.S. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3153 (1984).
- Seed, B. *Nucleic Acids Res.* **11**, 2427 (1983).
- Goldfarb, M., Shimizu, K., Peruchio, M. & Wigler, M. *Nature* **296**, 404 (1982).
- Chao, M. *et al.* *Cell* **32**, 483 (1983).
- Wright, S. *et al.* *Nature* **305**, 333 (1983).
- Diesseroth, A. *et al.* *Nature* **285**, 36 (1980).
- Charnay, P. *et al.* *Cell* **38**, 251 (1984).
- Palmiter, R. *Cell* **41**, 343 (1985).
- Scholnick, S.B., Morgan, B.A. & Hirsh, J. *Cell* **34**, 37 (1983).
- Spradling, A.C. & Rubin, G.M. *Cell* **34**, 47 (1983).
- Goldberg, D.A., Posakony, J.W. & Maniatis, T. *Cell* **34**, 59 (1983).
- Chada, K., Magram, J., Raphael, K., Radice, G., Lacy, E. & Costantini, F. *Nature* **314**, 377 (1985).
- Magram, J., Chada, K. & Costantini, F. *Nature* **315**, 388 (1985).
- Townes, T.M. *et al.* *EMBO J.* **4**, 1715 (1985).
- Hinnen, A., Hicks, J.B. & Fink, G.R. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4951 (1979).
- Scherer, S. & Davis, R.W. *Proc. natn. Acad. Sci. U.S.A.* **76**, 4951 (1979).
- Rothstein, R. *Methods in Enzymology* **101**, 202 (1983).
- Joyner, A. *et al.* *Nature* **305**, 556 (1983).
- Williams, D.A. *et al.* *Nature* **310**, 476 (1984).
- Anderson, W.F. *Science* **226**, 401 (1984).

Tom Maniatis is in the Department of Biochemistry and Molecular Biology, Harvard University, Cambridge, Massachusetts 02138, USA.

Erratum

In the article by Shen-su Sun "Ocean islands — plums or plumes?" (11 July, page 104) the last sentence on page 105 should have read: "It is conceivable that a large plum within the slowly convecting mantle might be sufficiently robust to turn into a plume . . ."

Identical pentapeptides with different backbones

SIR—In a comment on a recent *News and Views* article by Harrison¹, Fasman² maintains that we drew a false conclusion from the finding of identical pentapeptides with different backbone configurations³. He also gives the impression that his algorithm is capable of correctly predicting the secondary structure of such identical pentapeptides, as it takes into account the different neighbouring sequences on both their amino and carboxyl termini.

Contrary to Fasman's point of view, we believe that the strictly hierarchical approach of first predicting secondary structure segments, then tertiary structure by interaction of those segments, cannot succeed. There is a strong feedback component from the tertiary structure. This is clearly demonstrated in the case of bovine and porcine pancreatic phospholipases^{4,5}. Here, a segment of 13 amino acids differing only in one position (Val/Phe) assumes quite different conformations in the two closely related enzymes:

Bovine: ...LDCKVLVDNPYT...

Porcine: ...LDCKFLVDNPYT...

while the flanking sequences have the same conformation in both structures.

In conclusion, it has never been our intent to discredit predictive schemes but to point out their limitations⁶ and caution against overinterpretations. Hopefully, awareness of the difficulties of protein structure prediction will lead to real progress.

WOLFGANG KABSCH
CHRIS SANDER

Max-Planck Institut für
medizinische Forschung,
6900 Heidelberg,
Jahnstrasse 29, FRG

1. Harrison, S. *Nature* **313**, 736-737 (1985).
2. Fasman, G.D. *Nature* **316**, 22 (1985).
3. Kabsch, W. & Sander, C. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1075-1078 (1984).
4. Dijkstra, B.W., Weijer, W.J. & Wierenga, R.K. *FEBS Lett.* **164**, 25-27 (1983).
5. Lesk, A.M. *BioEssays* **2**, 213-214 (1985).
6. Kabsch, W. & Sander, C. *FEBS Lett.* **155**, 179-182 (1983).

Haldane's rule OK, but not always

SIR—In a recent article¹, Haldane's rule "When in the F_1 progeny of two animal races one sex is absent, rare, or sterile, that sex is the heterozygous (heterogametic) sex"², was lauded as surviving, presumably unblemished. It has been confirmed for mammals, birds, butterflies and snakes¹ and insects³. Nonetheless, a good rule need not be without exception.

In North America, certain populations of the mole salamander (genus *Ambystoma*), the female of which is the heterogametic (WZ) and the male the homogametic (ZZ) sex^{4,5}, exist as combinations of sympatric species and their hybrids. These complexes consist of naturally occurring diploid, triploid and tetraploid

individuals and such populations have been labelled unisexual, for in all those examined (*A. laterale-jeffersonianum*⁶, *A. jeffersonianum-texanum*⁷, *A. laterale-texanum*^{8,9}, *A. nothogenes*¹⁰) males have either yet to be found, or at most form 5% of the population. Of the few reported males, three have been triploid and the rest diploid. Laboratory reared offspring from eggs of diploid and polyploid females have all been female⁶. The rarity of males in these hybrid complexes contrasts strongly with the normal 1:1 sex ratio or even preponderance of males in non-hybrid ambystomid populations¹¹.

The rarity of males has prompted some authors to propose a parthenogenetic^{7,8} or gynogenetic⁶ mode of reproduction but we have shown that at least on Pelee Island (Lake Erie, Ontario), sperm is needed and incorporated into gametes of female *A. laterale-texanum*. Unreduced eggs of these hybrid females combining with haploid sperm produce the observed triploid and tetraploid animals. Unreduced eggs in a "dominant W" sex determination system will lead to mostly females^{4,5} as found, but diploid females should produce males in a minimum of 3F:1M ratio.

In hybrid *Ambystoma*, males, the rare or absent sex, are homogametic. Although there is reduced hatching success in eggs from hybrid females^{12,6}, triploid and even tetraploid hybrid females are fertile. In this early stage of our investigations, we conjecture that Z-containing gametes from diploid (WZ), Z or ZZ gametes from triploid (WWZ, WZZ) and Z, ZZ, or ZZZ gametes from tetraploid (WWWZ, WWZZ, WZZZ) females are rarely produced, are not viable, or that most larvae homozygous for Z and therefore destined to be males, do not survive. The virtual absence of polyploid males is explainable as a consequence or unreduced eggs from the W carrying females, but the number of diploid male hybrids is far lower than expected. It appears that in ambystomid salamanders, males may be innately more sensitive to the effects of interspecific hybridization than are the females, an explanation rejected by Haldane himself.

LAWRENCE E. LICHT

Department of Biology,
York University,
Toronto, Ontario,
Canada M3J 1P3

JAMES P. BOGART

Department of Zoology,
University of Guelph,
Guelph, Ontario,
Canada N1G 2W1

1. Jones, J.S. & Barton, N. *Nature* **314**, 668-669 (1985).
2. Haldane, J.B.S. *J. Genet.* **12**, 101-109 (1922).
3. Coyne, J.A. *Nature* **314**, 736-738 (1985).
4. Humphrey, R.R. *Am. J. Anat.* **76**, 33-66 (1945).
5. Seesions, S.K. *Chromosoma* **84**, 599-621 (1982).
6. Uzzell, T.M. Jr *Copeia* **1964**, 257-300 (1964).
7. Morris, M.A. & Brandon, R.A. *Copeia* **1984**, 324-337 (1984).
8. Downs, F.L. *Occas. Pap. Mus. Zool. Univ. Mich.* **685**, 1-36 (1978).

9. Bogart, J.P., Licht, L.E., Oldham, M.S. & Darbyshire, S.J. *Can. J. Zool.* **63**, 340-347 (1985).
10. Krauss, F. *Occas. Pap. Mus. Zool. Univ. Mich.* **709**, 1-24 (1985).
11. Husting, E.L. *Copeia* **1965**, 352-362 (1965).
12. Piersol, W.A. *Am. Nat.* **44**, 732-738 (1910).

Dating of craters and extinctions

SIR—Recently Weissman¹ discussed the topic of periodicity in the geologic record of extinctions and craters. The weakest correlation in his work is the late Ordovician extinction, for which his Fig. 1 shows no craters ≥ 10 km and only one > 5 km in diameter. On or about the same date I reported² the finding of iridium in breccias of Jephtha Knob, Kentucky and suggested that the ~ 2 km feature could represent the central peaks of a larger (~ 20 km?) complex crater. If this latter is true, because Jephtha Knob has intensely deformed Upper Ordovician rocks beneath undeformed Silurian cap-rocks, the correlations on Weissman's Fig. 1 would seem much better. Note also that the Glasford Disturbance, Illinois³, similar to Jephtha Knob but completely buried, is of the same age.

Many cryptoexplosion structures are poorly dated, but where their ages can be determined they should be included in the arguments for or against periodicity. Unfortunately, all in Grieve's⁴ list probably fail in size or in dating, although it is not clear which, if any cryptoexplosion structures are included in Weissman's Fig. 1.

C. RONALD SEEGER

Department of Geography and Geology,
Western Kentucky University,
Bowling Green,
Kentucky 42101, USA

1. Weissman, P.R. *Nature* **314**, 517-518 (1985).
2. Seeger, C.R., Asaro, F., Michel, H., Alvarez, W. & Alvarez, L. *Lunar planet. Science* **16**, 757-758 (1985).
3. Buschbach, T.C. & Ryan, R. *Bull. Am. Ass. Petrol. Geol.* **47**, 2015-2022 (1963).
4. Grieve, R.A.F. *Geol. Soc. Am. spec. Pap.* **190**, 25-37 (1982).

Icosahedral packaging of spheres

SIR—I wish to question the accuracy of the statement in A.L. Mackay's excellent letter (*Nature* **315**, 636; 1985) that the central space of icosahedrally packed spheres will accommodate a sphere with 95% of the radius of the others.

If the outer spheres are of unit diameter, then their centres form an icosahedron whose side is unity. The circumradius of that icosahedron is $5^{1/4} \cdot \tau^{1/2} / 2$ where τ is the Golden mean ($(5^{1/2} + 1)/2 = 1.618033989$). Thus the circumradius is 0.951056517. Since the outer spheres have a radius of 0.5 then the inner sphere has a radius of $(0.951056517 - 0.5) = 0.451056517$. This gives it a diameter of 0.902113034 and this is 90.2% of the outer spheres'.

It is of interest that the 12 spheres of the outer shell can be close packed in a 'dyma-

xion' array (Buckminster Fuller) and all touch an inner sphere of the same radius.

The point is not trivial, for I have been constructing models of quarks made up of infinite dimensional spheres of unit radius (each has a content then of 47.123789). Twelve of these in icosahedral array surround a central infinite sphere with 0.734 of the three-dimensional content of the outer spheres. This results in a content, taken over infinite dimensions, of 600.08 ($600 + (1/12)$). This allows the central space to be a four-dimensional 600-vertexed regular polytope. A quark is proposed as having an infinite kernel of 47.123789 giving 647.123789 for the u and d quarks. In a proton only a three-dimensional icosahedron is present in the kernel giving a content of $3 \times (600.08 + 12) = 1.836.24$.

In the neutron each icosahedral array changes to a dymaxion array (making it unstable) surrounding another sphere giving $3 \times (600.8 + 13) = 1.839.24$.

Some corroboration of this approach can be found in the examination of the content of different dimensional spheres, which rise to a maximum at five dimensions from unity at zero dimensions and reach unity again at 12.76 'fractal' dimensions¹.

BRIAN E. HAGAN

Queen Elizabeth II Research Institute,
University of Sydney and University
of New South Wales, Sydney, Australia

1. Hagan, B.E. *Appl. Phys. Commun.* 4, 161-177, (1984).

The distribution of charges in classical electrostatics

SIR—Berezin¹ has drawn attention to the "surprising" fact that the minimum-energy configuration of a system of N equal electric point charges q , constrained to lie on a disk of radius R , is in general not the arrangement with all the charges on the periphery with equal spacings (configuration A).

Berezin goes on to say that his result may have a bearing on truly three-dimensional geometries, but it can readily be proved that for a system of point charges confined to a spherical region, the minimum-energy configuration cannot conceivably have a charge at the centre.

Such a configuration would contradict Earnshaw's theorem² that a charge (say, at the centre of a sphere) cannot be held in stable equilibrium by the electrostatic field of all the other charges. This does not preclude Berezin's configuration B in the case of a disk, since a charge at its centre may be in stable equilibrium with respect to variations only in the plane of the disk.

Berezin's statement¹, that for N charges on the disk, one charge will be displaced to its centre (configuration B) for $N \geq 12$ is not valid.

With increasing N , the charges appear to arrange themselves in more and more complex patterns of concentric rings, with equal spacings on each ring. This problem was investigated by calculating numerical-

ly the total energy for various numbers of equally spaced charges on each of several such rings as a function of the radii of the rings, with a possible charge at the centre. The results are insensitive to the precise angular coordinates of charges.

With units in which $q = R = 1$ configuration B is that of minimum energy only for $12 \leq N \leq 16$. For $N = 17$, there is a lower-energy configuration with two charges at radius $r = 0.31$, the other 15 charges being at $r = 1$; its energy is $W = 133.8$, which is lower than the value $W = 134.1$ for configuration B.

When $N = 29$, for example, the lowest-energy configuration has six charges at $r = 0.49$, the remaining 23 being at $r = 1$ ($W = 444.5$, compared with $W = 459.4$ for configuration B). But when $N = 30$, the minimum-energy configuration has one charge at the centre, six at $r = 0.55$ and 23 at $r = 1$ ($W = 479.1$, compared with $W = 496.5$ for configuration B).

This trend suggests, as one might expect, that the charges tend to spread over the disk, though most of them always remain at the periphery. In fact, the continuum limit of an electrified conducting disk is known³. If the total charge is Q , the charge density at radius r is $\sigma(r) = Q(2\pi R(R^2 - r^2)^{-1/2})^{-1}$.

N.M. QUEEN

Department of Mathematics,
University of Birmingham,
PO Box 363,
Birmingham B15 2TT, UK

1. Berezin, A.A. *Nature* 315, 104 (1985).
2. Panofsky, W.K.H. & Phillips, M. *Classical Electricity and Magnetism*, 2nd edn, 26 (Addison-Wesley, Reading, Massachusetts, 1962).
3. Arfken, G. *Mathematical Methods for Physicists*, 2nd edn, 109 (Academic, New York, 1970).

SIR—Berezin¹ considers how N charges would distribute themselves over the surface of a disk and seems surprised to find that, for $N \geq 12$, they do not all lie around the circumference. But in the limit of very large N , they would surely be distributed with surface density $\propto (a^2 - r^2)^{-1/2}$, where a is the disk radius (for $r < a$), this being the charge distribution over a conducting disk given by simple electrostatics. The charges would actually group themselves into an N -dependent lattice pattern: it would be interesting (and perhaps relevant to snowflake growth, for instance) to calculate what these configurations are.

Classical electrostatics does *not* (as Berezin implies) tell us that the charge lies only at the edge of a two-dimensional conductor. (The charge density does, however, tend to infinity at the edge.) In three dimensions, the charge does lie on the surface and a Berezin-type calculation would confirm this rather than requiring "modification of usual theorems of electrostatic stability". This is readily seen for the case of a conducting sphere or ellipsoid. Suppose that the surface were charged and that one extra test charge were at the centre. If the surface charges were held fixed, the test charge would feel

no force even if it were displaced from the centre; however, the surface charges distribution would adjust (being repelled by the test charge), and the adjusted distribution would attract the test charge towards the surface, implying that no charge could remain stably in the interior.

MARTIN REES

Institute of Astronomy,
Madingley Road,
Cambridge CB3 0HA, UK

1. Berezin, A.A. *Nature* 315, 104 (1985).

SIR—My thanks to those who wrote comments on my recent letter 'An unexpected result in classical electrostatics'¹. The paradox is, indeed, resolved, if the circle with point charges is treated as a thin disk embedded in three-dimensional space and the central charge is actually at the centre of one of the flat surfaces of the disk. There will be, of course, no critical N for the 'charge ejection' to the centre of the case of a three-dimensional *empty* sphere. The problem of the gradual build-up of concentric spherical layers of point charges will, however, remain for the sphere containing a uniformly spread fixed positive charge (a static version of J. J. Thomson's atomic model).

This discussion has also revived an interest in the old but still largely unresolved problem of how to find the equilibrium configuration of N equal point charges placed on the surface of a sphere. Despite a number of valuable contributions²⁻⁵, its solution is known only for some special values of N and the general algorithm valid for any integer N is yet to be found. Even for the 'simple' cases of $N=4, 6, 8, 12$ and 20 (platonic polyhedrons) there is a very interesting curiosity which seems to contradict commonsense. While the tetrahedron ($N=4$), octahedron ($N=6$) and icosahedron ($N=12$) do indeed provide the minimum energy configuration, the cube ($N=8$) and dodecahedron ($N=20$) do not^{2,3,5}. It, perhaps, may be attributed to the fact that tetrahedron, octahedron and icosahedron all have 'rigid' triangle faces, while the cube and dodecahedron have 'soft' deformable faces (square and pentagon respectively).

Numerous other possibilities also remain open for study; for example, N non-equal charges ($Q, 2Q, 3Q \dots$), an unequal number of positive and negative charges confined to two close concentric spheres, point charges on non-spherical surfaces, and so on.

ALEXANDER A. BEREZIN
Department of Engineering Physics,
McMaster University,
Hamilton, Ontario,
Canada L8S 4M1

1. Berezin, A.A. *Nature* 315, 104 (1985). Comments: 315, 635; 316, 301-302 (1985).
2. Whyte, L.L. *Am. math. Monthly* 59, 606 (1952).
3. Leech, J. *Math. Gaz.* 41, 81 (1957).
4. Goldberg, M. *Math. Comput.* 23, 785 (1969).
5. Melnyk, T.W., Knop, D. & Smith, W.R. *Can. J. Chem.* 55, 1745 (1977).

The rough and the smooth

Sydney Brenner

The Transforming Principle: Discovering that Genes are Made of DNA.

By Maclyn McCarty.

W. W. Norton: 1985. Pp.252. \$14.95, £14.95.

FOR most young molecular biologists, the history of their subject is divided into two epochs: the last two years and everything else before that. The present and very recent past are perceived in sharp detail but the rest is swathed in a legendary mist where Crick, Watson, Mendel, Darwin — perhaps even Aristotle — coexist as uneasy contemporaries. It would not surprise me to find that most graduate students have not heard of Avery, MacLeod and McCarty or of their discovery that the transforming principle of the pneumococcus was DNA. The general ignorance of our times might easily consign this book, with the wonderful title *The Transforming Principle*, to the "Religion and Occult" section of a bookshop, as once I found Levi-Strauss's *The Raw and the Cooked* under "Recipes".

The book is Maclyn McCarty's account of his own life and, at the same time, the story, told from first-hand knowledge, of the discovery that genes are made of DNA (or, as it was seen then, that inherited characters could be introduced into bacteria by DNA). It is a story that is plainly and simply told, with none of the self-centred pomposity of so many autobiographies, and with an objectivity that increases its interest and fascination. McCarty came upon the scene quite late, and departed soon after the publication of the important paper; he seems to write about the work from the stance of a less involved observer but is nonetheless decisive and clear about his personal contribution. In addition, he has produced an interesting picture of the old Rockefeller Institute and an evocation of what it was like to be a scientist 50 years ago.

The sequence of events can be briefly recounted. Avery, like many of his contemporaries, had a lifelong interest in the pneumococcus, a great killer of pre-antibiotic days. In 1917, together with Dochez, he discovered that culture fluids of pneumococci contained a specific soluble substance (SSS) which could be typed with antisera. More or less concurrently it became clear that SSS was derived from the capsule of the bacterium and that virulence depended on the cap-

sule. Cultures of pneumococci underwent "dissociation" to throw off avirulent strains: on agar plates these gave "rough" (R) colonies, as opposed to "smooth" (S) colonies of the virulent, encapsulated strains. In 1926, Heidelberger showed that SSS, and hence the capsule, was a polysaccharide, and, two years later, Dubos was able to enrich for a bacterium which produced an enzyme that degraded SSS type III which was to play a significant role later.

It was in the course of study of the smooth and rough variation that Griffith discovered transformation. In a 46-page paper published in 1928, he reported that mice injected with an avirulent, rough strain, derived from one type, together with a heat-killed smooth strain of a different type, succumbed to infection; remarkably, the virulent organism isolated was of the type of the dead smooth strain.

Thus, type I rough strains could be "transformed" into type II or III, and type II into type I or III by using the appropriate heat-killed strains. Griffith tried to get transformation *in vitro*, but failed. His interpretation of the results can be judged from the following remark in his paper:

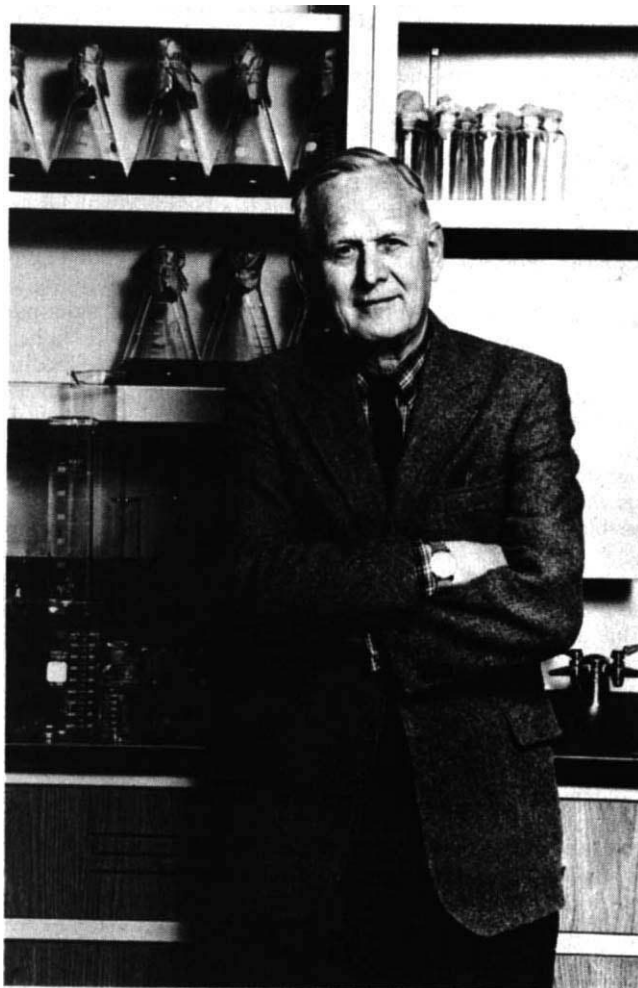
When the R form of either type is furnished under suitable experimental conditions with a mass of the S form of the other type, it appears to use that antigen as a pabulum to build up a similar antigen and thus develop into an S strain of that type.

There was no genetics involved here, and his idea was very much that the antigen acted as a "primer" and itself determined what antigen was subsequently made.

The *in vivo* research was quickly confirmed by other workers, in particular by Dawson who was working in Avery's laboratory on S-R interconversion. His 1930 paper also reported the failure to produce transformation *in vitro*, but, together with Sia, he was able to announce success a year later, after he had left the Rockefeller Institute. Alloway continued the experiments in Avery's laboratory and by 1933 had found a way of making extracts with consistent transforming activity. None of these papers had Avery's name on them but there is evidence that he followed the work closely.

The hero of the tale is, without doubt, Colin MacLeod, who entered Avery's laboratory in 1934 to work on transformation. He selected a stable rough strain, R36, which showed no reversion and was efficiently transformed. He used Dubos's enzyme to prove that transformation by type III could not be due to the polysaccharide; this method of using specific enzymes to assay the chemical nature of the transforming principle was to be applied over and over again, and indeed was important to the final proof that it was DNA. MacLeod aroused Avery's interest in the work again, and research proceeded for three years when it was dropped, and MacLeod went to work on other aspects of the pneumococcus. It seems that he needed to push out publications so as to ensure his future career. Times have not changed very much so far as this familiar problem is concerned, but what has changed is that nobody else entered the field during the three-year hiatus.

In 1940 the research was taken up once more, but MacLeod left the Rockefeller Institute in 1941. Nearly all the subsequent work was done by McCarty, who joined Avery's laboratory in the same year, and in the book he



Maclyn McCarty — "This is it, at long last".

gives a first-hand account of all the experiments which led to the conclusion "that a nucleic acid of the deoxyribose type is the fundamental unit of the transforming principle of *Pneumococcus* Type III". The famous paper appeared on 4 February 1944 in the *Journal of Experimental Medicine*; on the reprint he sent to his mother, McCarty wrote "This is it, at long last".

Gunther Stent has tried to explain why this result had so little impact on contemporary genetics by arguing that it was a "premature" discovery. McCarty does not accept this. He believes that it made possible the development of molecular genetics but that "it required further biological, chemical, and structural development before it could be manipulated by geneticists". Bacteriology developed quite separately from the rest of biology and was connected strongly with medicine. All of the participants were trained in medicine and were interested in infectious disease; they did not believe for quite some time that they were in fact doing genetical research, although they were clearly conscious of the importance of the result once they got on the right track. I think Stent is right in saying that the field was not ready to assimilate this result. There was no way at that time that anybody could have conceived how to go from DNA to polysaccharide. The role of proteins was not understood, the one-gene-one-enzyme hypothesis had not been formulated, and haploid genetics in organisms with no visible chromosomes had yet to be invented. Indeed there were many people who believed that bacteria did not have genes because Mendelian experiments could not be performed with them, and it would have been an impossible leap for geneticists to view the transformation experiments as a kind of genetic cross. In 1952, Hershey and Chase showed that when a bacterial virus infects a bacterium only the DNA enters the cell while most of the protein is left on the outside. Although this experiment was not as "clean" as the transformation experiments, it had a greater impact precisely because of the intervening conceptual developments that had taken place in fields such as biochemical genetics.

Did this work deserve a Nobel Prize? Of course it did, and probably more so than many of the ones given since. But as McCarty suggests, the Committee was more careful than wise and let it go by. Thus the story ends on a somewhat sad note, and leads the reader to contemplate the ruthlessness of the process of scientific research, and why, as one grows older, one comes to value books such as this, not so much for their scholarship, but for the memories of men and the humanity they bring with them. □

Sydney Brenner is Director of the Medical Research Council's Laboratory of Molecular Biology, Hills Road, Cambridge CB2 2QH, UK.

Homage to Patagonia from the hard-nosed palaeontologist

Leonard Krishtalka

Bone Hunters in Patagonia: Narrative of the Expeditions. By J. B. Hatcher. *Ox Bow Press, PO Box 4045, Woodbridge, Connecticut 06525, USA: 1985. Pp.209. Hbk \$24; pbk \$12.95.*

FEW FIELD journals are classic, slow-paced thrillers. Darwin's *Voyage of the Beagle* is one. John Bell Hatcher's *Bone Hunters in Patagonia* is another. From March 1896 to September 1899, Hatcher, a renowned American palaeontologist and collector, led three year-long scientific expeditions to Patagonia on behalf of Princeton University. Like Humboldt and Darwin before him, Hatcher was one of a succession of illustrious naturalists lured to South America by its strange faunas, living and fossil. What drew Hatcher in particular were the exotic, extinct marsupial and placental mammals preserved in Patagonia's basaltic badlands and towering sea cliffs. The results of his work — the seven-volume *Reports of the Princeton University Expeditions to Patagonia* — appeared in 1903; four of the volumes described Hatcher's huge palaeontological collections and three equally massive ones the extant mammals, birds and plants. Buried in the first volume of the *Reports*, and rarely read, is Hatcher's fascinating "Narrative of the Expeditions", of which *Bone Hunters in Patagonia* is a facsimile reprint. As such, it is a rhapsodic yet blood-and-guts overture to an early scientific masterpiece and ranks with the South American sections of Darwin's *Voyage* in its mix of adventure, adversity and natural history.

Nine chapters of *Bone Hunters in Patagonia* describe the first expedition (March 1896 to July 1897), two deal with the second (November 1897 to November 1898) and one covers the third (December 1898 to September 1899). Throughout, Hatcher is a study in dualism. His travails with equipment, horses, food, disease, injury and Patagonian winters and terrain are, in gentlemanly fashion, understated. Twice he came close to death, once from a severe head wound, and again from a violent recurrence of inflammatory rheumatism, a lifelong disease. From between his lines emerges a "wild-west" loner and hard-nosed perfectionist of unshakable determination and courage who felt driven "beyond the limits of civilisation [to study nature] in her true form" (p. 169).

Hatcher bitterly resented the "fireside naturalist, who seldom goes beyond his private study" (p. 37), especially those "parasites" who gained palaeontological

repute by describing his collections. Rural Argentina too evoked his ethnocentric ire, from the cold, filthy villages to the overrated talents of the gauchos. But his view of Patagonia's natural history and geography is almost elegiac. Here, Hatcher's "love of nature" brought a focused eye and sensitive pen to the appearance, ecology and behaviour of myriad species of mammals, birds and lizards, and made him rue his "brutal" killing of these animals, whether for food or science. He was awed, almost seduced, by the grand contours of Patagonia's coastal, interior and Andean terrain, much of which Hatcher was the first to explore and map. His geographical discoveries included new mountains and rivers, and helped settle the Argentinian-Chilean boundary question.

Hatcher makes only passing reference to the geology and palaeontology of Patagonia, subjects he reserved for the scientific volumes but never completed. He does exult (deservedly so) over the wealth of Santacrucian (Miocene) fossil mammals he and his sole assistant, O.A. Peterson, collected at Killik Aike (1.5 tons) and Corriguen Aike (4.5 tons) and correctly challenges Florentino Ameghino's erroneous association of dinosaurs (Cretaceous) with the mammalian fauna from the *Pyrotherium* beds (Oligocene). He was unable to find these beds, however, and as this was the goal of the second and third expeditions, it was Hatcher's only failure, and one he constantly lamented. Nevertheless, his geological work helped refine the stratigraphy of Patagonia, and his voluminous collections of Patagonia's extant and fossil animals and plants — many of them new species — allowed an unprecedented comparison of faunal composition and evolution in the Americas. Curiously, Hatcher does not mention his notion of a single southern landmass, an idea he hoped to explore during a fourth expedition in 1903 while employed at the Carnegie Museum, Pittsburgh.

This Patagonian narrative should stoke the spirits of all palaeontologists, geologists and field biologists, and will coax the "fireside naturalist" away from the fire. Its appearance now as *Bone Hunters in Patagonia* is especially timely, as it can be read alongside G. G. Simpson's recent (and excellent) history of palaeontological explorations in South America (*Discoverers of the Lost World*; Yale University Press, 1984). There are two epilogues to the narrative: both are sad and ironic. Five years after surviving the rigours of the expeditions, Hatcher died of typhoid fever contracted from Pittsburgh's untreated drinking water. And in 1985, Princeton University gave away its palaeontological collections, including Hatcher's hard-won lode of Patagonian fossils. □

Leonard Krishtalka is Associate Curator in the Section of Vertebrate Fossils, Carnegie Museum of Natural History, 4400 Forbes Avenue, Pittsburgh, Pennsylvania 15213, USA.

Three dimensions to theory in chemistry

Jennifer Green

Orbital Interactions in Chemistry. By T.A. Albright, J.K. Burdett and M.H. Whangbo. Wiley: 1985. Pp.447. \$49.95, £51.10.

CHEMISTS are very receptive to pictorial representations of their science, and are well-used to thinking in three dimensions. Albright, Burdett and Whangbo take full advantage of this ability in their book, which contains a myriad of graphic representations of orbital interactions in molecules. Their aim is to describe and examine the electronic structure of covalent molecules and solids with a view to understanding their existence, structure and, in some cases, their reaction pathways. The strength of the semi-empirical approach on which such orbital descriptions are based, and the "cartoons" in which the results are communicated, is their accessibility to the non-mathematical chemist, who is thereby provided with a set of theoretical tools for prediction, as well as rationalization, of structure and reactivity.

In a carefully planned first section the quantum mechanical and group theoretical basis of this semi-empirical approach is clearly expounded. Within a framework of LCAO-MO theory, symmetry, overlap, electronegativity and perturbation arguments are used to establish the orbital structure of simple cases. Here many motifs are outlined which recur, with elaborations, throughout the book. The implications of various occupancies of these prototype orbital structures are discussed; exchange energy, first and second Jahn-Teller effects, the Fock operator and configuration interaction are brought into play.

The second section consists of an exploration of main group covalent molecules using the technique of building up the orbital structure of larger molecules from that of smaller fragments, while maintaining a delocalized picture of the orbitals. Where relevant, the relation to a localized bond picture is made clear. The third section is a similar treatment of transition metal complexes. The extensive range of molecules discussed, from the simplest organics to metal clusters, bears witness to the power of the technique.

New editions

- *The Meaning of Fossils: Episodes in the History of Palaeontology*, 2nd Edn, by Martin J.S. Rudwick. Publisher is University of Chicago Press, price is pbk \$11.95, £11.25.
- *Flame and Combustion*, 2nd Edn, by J.A. Barnard and J.N. Bradley. Publisher is Chapman and Hall, price is pbk £12.95, \$27; hbk £25, \$55.

However, as the density of examples increases the clarity of the explanations declines. The emphasis on coverage, rather than exemplification, makes the text less accessible than it might have been to the undergraduate, but more useful to research chemists who will appreciate having this large body of work brought together within a single volume.

Throughout, the unity of approach is impressive for a book written by three authors, and can be attributed in part to the influence, acknowledged in the preface, of Roald Hoffmann (who also contributed the foreword). The style of writing is less even but at all points the illustrations help in understanding the text and vice versa. Particularly welcome is a chapter on solids, which attempts the Herculean task of translating the language of

solid state physics for the benefit of the chemist with only a moderate acquaintance with molecular orbital theory. For this to be entirely successful a complete book would be needed, but this short account is very helpful.

The book gives a rather bland impression of the periodic table; perhaps the blend of overlap and symmetry arguments could have been seasoned by more emphasis on the periodic trends of orbital energies and separations, which contribute to the uniqueness of each element's chemistry. However, as a compilation and clear exposition of the work of an influential school of theoretical chemistry it will be extremely useful. □

Jennifer Green is a Fellow of St Hugh's College, Oxford OX2 6LE, UK.

A wealth of viruses

D.J. McCance

Virology. Edited by Bernard N. Fields. Raven: 1985. Pp.1,614. \$176.50.

AT 4.75 kilograms there is no disputing the publisher's intention that this tome should be a reference book; no one would consider carrying it around for light reading. With the title *Virology* one would expect a broadly based text and indeed this is the case, except that only mammalian viruses of medical importance are discussed (plant viruses are relied on heavily in the sections on viral structure however).

This is the first attempt to cover mammalian virology comprehensively since 1974 when Fenner *et al.*'s *The Biology of Animal Viruses* appeared, and the increase in size over that book reflects the advances of the past ten years in our knowledge of virus structure, genetics, replication and pathogenesis. The dangers of a volume such as this, of course, are that with so many contributors (here 64) it takes time to come to press and because of constraints on space the authors may be tempted into superficiality. Both pitfalls have been avoided: the various chapters are rich in detail and, judging by the dates of papers in the reference lists, the editor managed to cajole his authors into meeting their deadlines.

There are really two parts to the book, the first third of it dealing with the general principles of medical virology and the rest with individual viruses or virus groups. In this second part, the molecular biology and pathogenesis of particular viruses are discussed in successive chapters, thus providing easy access to information of interest to people of different backgrounds. With this format the book will be useful to a wide range of individuals, including undergraduates from related disciplines, molecular virologists and clinicians. The clinician should take comfort

in the fact that in general the molecular virology aspects are well illustrated with informative and digestible figures to help in comprehending the complexities of virus structure and replication.

The wide-ranging nature of the book gives the reader an interesting perspective on the state of virology during the past decade, which has seen the discovery, for example, of the human retroviruses, HTLV I-III, the rotaviruses causing gastroenteritis and the hepatitis viruses. It is also possible to discern the renewed emphasis placed on previously identified virus groups, often because of the inroads made through molecular biological techniques. A decade ago the pathogenesis of the genus papillomavirus would have only been mentioned. Now, given their association with some genital malignancies, they command half the chapter on the Papovaviridae, while hepatitis A and B, non-A, non-B and the Delta agent take three chapters. Unfortunately genetic engineering has not been as productive in other areas, such as viral chemotherapy, where progress has been much slower.

Overall the book succeeds in covering medical virology in depth, and with its helpful illustrations and extensive bibliographies will find wide use as a reference text. It is clear that although great progress has been made on the molecular side of the subject, our knowledge of the process of pathogenesis lags well behind. One can hope that the book's structure will encourage those from either molecular or clinical backgrounds to browse through both types of chapter, and result in more people becoming interested in pathogenesis at the molecular level. Such individuals are needed to utilize molecular techniques in investigating the pathogenesis of virus diseases and to produce effective treatment and preventive measures. □

D.J. McCance is Senior Lecturer in the Department of Microbiology, United Medical and Dental Schools of Guy's and St Thomas's Hospitals, Guy's Campus, London Bridge, London SE1 9RT, UK.

Science and medicine put right

D.J. Weatherall

Communicating in Science: Writing and Speaking. By Vernon Booth. Cambridge University Press: 1985. Pp.68. Pbk £3.95, \$6.95.

Research: How to Plan, Speak and Write About It. Edited by Clifford Hawkins and Marco Sorgi. Springer-Verlag: 1985. Pp.184. Pbk DM32, £8.75.

BROWSING through these books evoked painful memories. In the middle of the sleepless night before my first scientific presentation I discovered several mistakes in my slides. Panic stricken, I roused the night porter of the hotel in which I was staying and, after persuading him with difficulty that I was not putting the finishing touches to a bomb in my bedroom, borrowed a roll of electricians' tape with which to disguise the worst of the errors. When projected in a large lecture theatre the next morning the doctored slides appeared to be peppered with enormous black holes, an adornment which, though it caused a burst of whispered speculation, did little to clarify my garbled message, particularly because by then I was too far gone to see my notes. As I was helped from the podium I vowed that, if ever I ran a department, no youngster would be let loose on the public without rehearsal.

It is not surprising that scientists find it equally difficult to write. Many of us progress to our first research fellowships without having to write a decent piece of prose; as pointed out by George Pickering, the invention of the multiple-choice examination was the death blow to competent writing. Yet as the pressure to publish becomes greater, there is less time to correct or polish our efforts. Editing a multi-author textbook is a salutary experience. The only charitable explanation for some of the offerings that arrive from distinguished colleagues is that they were generated somewhere in the brainstem and transferred to the editor via a dictating machine and a word processor, without further human contact.

These two books aim to improve the situation. Not surprisingly, they are both much more successful when dealing with writing than with speaking. Vernon Booth's book, an extension of his excellent essay *Writing a Scientific Paper*, is a mine of helpful information on style and clarity, and could be read with profit by scientists of all ages. Of the new material, the short section on preparing a thesis will be invaluable to research students, but the chapter addressed to North Americans is slightly irritating. Though apologetic, Dr Booth makes it clear that he does not approve of a lot of American scientific writing. But are Americans really more

guilty than anyone else about the use of "stacked modifiers" and unnecessarily long words, and the misuse of the future tense? And why should Americans be expected to adopt British usage? A rich language is constantly changing; the more ghastly neologisms, from which ever side of the Atlantic they emanate, are unlikely to survive.

Hawkins and Sorgi's book is a multi-author effort directed primarily at young clinicians. As well as guiding them on how to write and speak, it tells them why they should do research and, perhaps rather ambitiously, how to do it. David Heath, in the introductory chapter, suggests that most doctors should spend two to three years of training in full-time research, "preferably in a laboratory". This, he says, will shape their critical faculties and turn them into better doctors, an attitude which is deeply entrenched in academic

a mediaeval magic lantern which is still in use in one of Britain's Royal Colleges, or have had to suffer the sensory deprivation caused by lecturing, slideless and in complete darkness, during a power cut in Athens, will any real feeling of confidence in communication be engendered in young scientists. And too much confidence can be worse than abject terror. Distinguished speakers, their over-exposed grey diazo slides matching their hair, and too blasé to put their thoughts into order, are much less interesting than panic-stricken beginners. Audiences thrive on tension, and eccentricity. Lectures by the eminent Cambridge physiologist who chewed his tie while he talked, often to the extent of gagging and having to start again, or the London physician who let his glass eye drop out if he sensed that his audience was flagging, were riveting experiences. It would be a



medicine. But is it entirely correct? While most doctors worth their salt will find rewarding clinical problems to explore, it is only a minority that have the ability and, even more important, the motivation to spend several years in a laboratory. They should not feel that it is essential to drop their clinical work simply to further their careers, a formula for frustration and poor science.

Hawkins and Sorgi cover a much wider field than Booth. There are informative chapters on statistics, reference filing and illustration, and appendices on thesis-writing, useless words, the anatomy of the dictating machine, and, again, American usage. All aspiring medical research workers should have access to this beautifully written and very practical book.

Although both of the books are excellent guides to scientific writing, they left me wondering how far it is possible to define — let alone teach — the skills of talking about science. Certainly there is no substitute for exhaustive rehearsal in front of a ruthless group of colleagues, and bitter experience. Only when they have seen their slides ground to dust by an aggressive Chinese slide projector, or tossed into the air like toast from a toaster by

pitiful, in our well-intentioned efforts to encourage competence in teaching and scientific presentation, we bred too much uniformity of style. Scientific research is tremendous fun; hearing about it should be equally diverting. □

D.J. Weatherall is Nuffield Professor of Clinical Medicine at the University of Oxford, John Radcliffe Hospital, Headington, Oxford OX3 9DU, UK.

Fresh references

- *Genetic Engineering and Biotechnology Yearbook 1985*, edited by Alan G. Walton and Sharon K. Hammer. The Yearbook is an updating and expansion of the previous edition of 1983, and incorporates more information on the financial performance of biotechnology companies as well as details of their structure, personnel, research and products. Publisher is Elsevier, price is Dfl. 2,000, \$600.

- The first three volumes of the eighth and English edition of *Geigy Scientific Tables*, recently available in Britain from Farrand Press, 50 Ferry Street, Isle of Dogs, London E14 9DT. Like its predecessors, the eighth edition contains a wide range of data for clinicians and medical scientists in particular; unlike them, it is to appear in a number of volumes, and is not to be given away. The price (subsidized by Ciba-Geigy) is £12.50 per volume.

Production and destination of British civil plutonium

K. W. J. Barnham, D. Hart, J. Nelson and R. A. Stevens

The amount of plutonium produced by the Magnox reactors belonging to the CEBG and SSEB is estimated using three different methods which give similar results for total plutonium production. The difference between this total and the UK civil plutonium inventory is 6.3 ± 0.8 tonne. This balance was apparently sent to the United States in exchange for fissile material for UK military requirements. The US destinations published by the UK government appear to accommodate significantly less plutonium.

WE believe that if the international non-proliferation regime is to be strengthened, all links between civil and military nuclear programmes should be broken. If such links existed in the past the details should be clarified, effective safeguards introduced to prevent re-occurrence and information made available to allow compliance with the safeguards to be monitored. Within the framework of the Non-Proliferation Treaty (NPT) these injunctions clearly apply to the three nuclear powers which are signatories of the treaty, Britain, the Soviet Union and the United States.

Accordingly, we have set out to determine, for the case of the United Kingdom, how much plutonium has been produced in its civil reactors. Because much of the information required for such an investigation is not available, we have used three methods to estimate the total plutonium production. We believe our calculations are an interesting demonstration of the precision with which the production of fissile material such as plutonium-239 may be inferred from published data about reactor operations.

The British situation is complicated by the arrangement sanctioned by the mutual defence agreements with the United States under the terms of which¹ plutonium from the British civil programme has been transferred to the United States in exchange for highly enriched uranium and tritium required for the British military programme. British government spokesmen have stated² that no plutonium from the British civil reactors operated by the Central Electricity Generating Board (CEGB) and the South of Scotland Electricity Board (SSEB) has been used for military purposes in Britain and that none of the plutonium transferred to the United States has been so used. Our calculations, however, lead us to conclude that the civil uses for British plutonium in the United States do not fully account for the missing material. We have also encountered what we believe are serious shortcomings in the procedures for plutonium accountancy in Britain.

An earlier version of this paper was first submitted for publication in June 1984 and presented at the Sizewell Inquiry in October 1984³. The CEGB responded⁴ with a number of criticisms, which they did not quantify on advice from the Department of Energy. We have investigated all these criticisms and incorporated some changes. These make only small differences to our numerical results.

Yield and burn-up

The British Magnox reactors with which we are concerned are graphite-moderated reactors using natural uranium fuel. A series of these reactors were built in Britain primarily for production of military plutonium. The first of these, at Calder Hall in Cumbria, was commissioned in 1956; others followed at Chapelcross in Scotland. We are not here concerned with the production of plutonium at these reactors, but at the civil reactors designed on similar principles but operated by the electricity utilities.

Our first objective is to calculate the total quantity of all plutonium isotopes produced per tonne of original fuel as a function of the total thermal energy generated by a tonne of fuel, called the burn-up, B , in units of MW-days per tonne (MWd/te). We represent plutonium isotope production by the function $G(B)$, in units of kilograms per tonne (kg/te). This can be obtained by the numerical solution of the equations⁵ giving as a function of irradiation time the concentrations of the principal isotopes of uranium (235 and 238) and of plutonium (239 to 242 inclusive). The burn-up, the energy released by fission of uranium 235 and 238 and of plutonium 239 and 241, is a by-product of these equations. We correct the energy release for radiative capture effects as in ref. 6. The cross-sections we use⁷ depend on the mean neutron temperature which we take to be the mean moderator temperature T_m , and on the proportion of epithermal neutrons in the total neutron spectrum governed by the parameter r . Resonance absorption and fast fission are allowed for by the adjustment of the uranium-238

cross-section using parameters for each reactor given by IAEA⁸. A detailed description of our calculations will be published elsewhere.

There are no direct tests we can make of our $G(B)$ on data for the CEGB and SSEB reactors. The only clear information published on plutonium production in Magnox reactors concerns the military reactors at Calder Hall. Figures by Tyror⁹ and Griggs and Harper¹⁰ illustrate plutonium production and isotopic composition as a function of burn-up for a typical point in the Calder Hall reactor core. At a moderator temperature appropriate to Calder Hall and with $r = 0.055$, we find our $G(B)$ reproduces these curves very closely. Our best fit to the Tyror curves for isotopic composition is obtained with a fast fission factor of 1.02, corrected for fast radiative captures, as used by Griggs and Harper¹⁰, rather than using 1.03, as quoted by the IAEA⁸.

The major difference in $G(B)$ between Calder Hall and the civil Magnox reactors arises from different moderator temperatures^{4,8}. For each civil reactor we use the $G(B)$ described above evaluated at the temperature appropriate to that reactor^{8,11}. The effect of changing other parameters is considered later.

Table 1 Comparison of isotopic ratios in spent fuel dispatched 1978–84 with predictions for discharges 1977–83

	Plutonium-239 Sum all plutonium isotopes	
	$G(B)$ predictions 1977–83	CEGB dispatch data 1978–84
Bradwell	0.730	0.744
Berkeley	0.740	0.721
Hinkley Point A	0.740	0.725
Trawsfynydd	0.719	0.717
Dungeness A	0.703	0.719
Sizewell	0.718	0.716
Oldbury	0.704	0.708
Wylfa	0.705	0.710
Mean	0.720	0.720

Table 2 Uranium fuel (tonnes) discharged from CEBG and SSEB power stations 1963–72

Fiscal year	Bradwell			Berkeley			Hinkley Pt			Trawsfynydd		Dungeness			Sizewell		Oldbury		Wylfa		Hunterston*	
	IME	SUB	BEST	IME	SUB	BEST	IME	SUB	BEST	SUB	BEST	IME	SUB	BEST	SUB	BEST	SUB	BEST	SUB	BEST	IME	BEST
63–64	44	40	44	45	49	45																
64–65	83	93	83	60	43	60																
65–66	133	120	133	188	147	188	10	63	10	33	33	6	63	6								
66–67	117	120	117	128	147	128	47	63	47	33	33	67	63	67	48	48					29	29
67–68	97	120	97	118	147	118	208	63	208	33	33	146	63	146	48	48					99	99
68–69		74	74		110	110	440†	408	408	73	73		178	178	212	212	44	44			178‡	178
69–70		94	94		118	118		174	174	170	170		173	173	149	149	91	91			122‡	122
70–71		93	106		116	118		100	105	155	142		169	182	142	167	110	118			73‡	73
71–72	81	68‡		106	104‡			18	13‡	52	65‡		158	145‡	160	135‡	136	128‡	16	14‡		42‡

* Calendar year basis, given by first year in first column.

† Based on rate for April–August 1968.

‡ Assuming proposed curve followed.

§ Ref. 18.

|| Ref. 15.

This $G(B)$ refers to a point within the reactor core. The spatial variation of neutron flux in the core gives rise to a spread in the irradiation of the fuel elements in any particular channel. We have investigated typical axial variations of burn-up for a Magnox reactor¹² and used such a variation to obtain a channel-averaged $G(B)$. As $G(B)$ is approximately linear over the appropriate range of burn-ups, averaging reduces $G(B)$ by only 1–2 per cent. Possible radial variations in burn-up have also been investigated. Though the neutron flux falls at large radius, we find that for sensible refuelling procedures in the steady state, radial averaging compensates in part for the effect of axial averaging. A comprehensive reactor-average of $G(B)$ is not feasible in the absence of detailed information on refuelling schemes. Henceforth we use a $G(B)$ that is channel-averaged only, noting that this will underestimate the real situation.

We have been unable to find detailed information against which to test our $G(B)$ for civil Magnox reactors apart from Fig. 5 of the sixth report of the Royal Commission on Environmental Protection (the "Flowers Report")¹³, which gives the rate of plutonium production for a "power" reactor. Our original interpretation of this unreferenced figure as typical of civil Magnox reactors has been criticized by the CEBG⁴, but our $G(B)$ does reproduce this curve well at a temperature within the range appropriate to civil Magnox reactors.

As an additional test, we have compared the isotopic ratios resulting from our calculations with the data provided by the CEBG on the isotopic composition of fuel dispatched from their Magnox stations for the six fiscal years 1978–84¹⁴. We believe that the isotopic composition of the fuel dispatched should approximate to that of the fuel discharged one year earlier. Table 1 shows the good agreement between our predicted isotopic ratios, calculated at the average of the discharge burn-ups for the appropriate years (as in Method A below) and the CEBG dispatch

data, giving confidence in our extrapolation of the Calder Hall fit.

Fuel discharges

Fuel discharges from 1971–72 onwards have been provided by the British government in response to parliamentary questions^{15–21}, but the government has refused to give information on fuel discharges for the 1960s¹⁵. The figures in Table 2 in the column headed "SUB" are derived by subtraction of the numbers of fuel elements discharged by certain dates given in various sources^{21–25}. For five stations, refuelling curves to mid-1968 are given in an Institute of Mechanical Engineers symposium (IME)²⁶ on the refuelling of gas-cooled reactors. Bearing in mind that the SUB data for 1965–68 are averaged over three years, the two sources are in reasonable agreement.

The refuelling policy adopted at Magnox stations in their early years was to follow an "ideal refuelling line"²⁶ with the total spent fuel discharged, ΣD_i , increasing linearly with "core-average" irradiation

$$B = (\Sigma E_i)/M \quad \text{MWd/te} \quad (1)$$

up to a predetermined maximum burn-up $B_{\max}$. Here ΣE_i is the total thermal energy generated (in MWd) and M is the total mass of uranium in the core. If this "ideal refuelling line" is followed, then all of the initial charge will have been discharged by the time the "core-average" burn-up reaches $B_{\max}$. In addition, if each D_i discharged is replaced by an equal amount of fresh fuel, then when $B_{\max}$ is reached the core will contain fuel with all burn-ups equally represented. This is the ideal steady-state situation.

The refuelling curves available²⁶ show that the Magnox stations fell behind the ideal refuelling line in the early years, but that in the late 1960s, strenuous efforts were made to increase refuelling rates until the ideal, or a line parallel to it, was achieved. As operating experience was acquired, $B_{\max}$ was increased, so that the

steady-state description only approximately represents the situation in the 1970s. The effects of such factors will be considered later.

Two of our models require E_i , the thermal energy generated. For CEBG stations we have obtained these by fiscal year from the CEBG²⁷ and parliamentary answers^{16,17} and for Hunterston A by calendar year up to 1982 from the SSEB¹⁸ and fiscal year subsequently^{19,20}.

The models

Given the fuel discharges of Table 2 and $G(B)$, only the burn-up at which the fuel was discharged is needed to calculate the plutonium production. The CEBG, however, have refused to provide average discharge burn-ups²⁸. In method A, we have taken discharge burn-ups from a number of sources²⁹. For other years we linearly interpolate between these published figures or between the earliest published figure and zero burn-up on starting up. We then calculate plutonium discharge using these discharge burn-ups, the fuel discharges of Table 2 and our channel-averaged $G(B)$. The totals to 31 March 1985 for each reactor are presented in Table 3.

Method B also uses the fuel discharge figures of Table 2, but attempts a more detailed calculation of burn-ups using figures on the thermal energy generated per year. We increment the core-average burn-up by E_i/M each year and determine the burn-up each batch would receive by mid-year. Discharged fuel is replaced by equal amounts of fresh fuel, the burn-up of which we increase by the core average in subsequent years. When all the initial charge is discharged, we then discharge the fuel loaded in the first year assuming a policy of "first in, first out". Using a computer program for the book keeping, we find that the burn-ups of the spent fuel discharged in the steady state are similar to but in general slightly lower than the discharge irradiations assumed in Method A. This result is expected as we calculate an average burn-up and in practice chan-

nels with higher than average burn-up will be preferentially discharged. As a result Method B probably underestimates plutonium discharged and overestimates plutonium in core. In fact the total plutonium in core at 31 March 1985 according to Method B is 9.6 te to be compared with the value of 9.5 te quoted to the nearest half-tonne in a parliamentary answer³³, suggesting that this can only be a small effect. We determine plutonium discharge at each burn-up from $G(B)$ and the size of the batches. The results obtained from Method B are presented in Table 3.

A further test of Method B is provided by the plutonium content of the fuel dispatched from CEGB stations in the years 1978-84¹⁴. Seven CEGB stations keep discharged fuel in cooling ponds where the fuel cannot remain indefinitely because the cladding would corrode. The average time between discharge and dispatch for the fuel in ref. 14 was 1.2 years. However this average probably includes Wylfa which has a dry store. We believe that, for the stations with cooling ponds, the total plutonium dispatched over a six year period should be similar to the total of the plutonium produced in a similar period starting one year earlier. We compare CEGB dispatch data with appropriate Method B production figures in Table 4. Note that the reactor-to-reactor variation given by Method B is similar to that in the CEGB data and overall our predictions are a 3.6% underestimate. The CEGB have refused to publish¹ totals of plutonium in the ponds at the start and the end of the six-year period which could discredit or confirm our calculations.

Method C uses the total thermal energy generated and does not use any figures for spent fuel discharged. The principle of the method is to assume that the "ideal refuelling line" was followed. If a linear rise to a certain maximum value of burn-up ($B_{\max}$) is assumed then the energy extracted from fuel in core on rise to steady state,

$$E_c = \int_0^{B_{\max}} B dD = MB_{\max}/2$$

and E_c is equal to E_d , the energy extracted from the fuel discharged in this period.

If in the steady state the amount of fuel D_s is discharged at the burn-up, $B_{\max}$ over a period of time in which thermal energy E_s is generated, then it is straightforward to show that to keep the steady state situation constant

$$D_s = E_s/B_{\max} \quad (2)$$

Therefore total thermal energy generated

$$E_T = E_c + E_d + E_s = MB_{\max} + D_s B_{\max} \quad (3)$$

Plutonium production can similarly be divided into three parts. Plutonium in core at start (and end) of steady state

Table 3 Plutonium discharge (te) by year (Method B) and totals (Methods A, B, C)

	BRADWELL	BERKELEY	HINKLEY POINT A	TRAWSFY- NYDD	DUNGENESS A	SIZEWELL	OLDBURY	WYLFA	HUNTERSTON A (CALENDAR YEARS)	
63-64	0.02	0.03								
64-65	0.09	0.06								
65-66	0.19	0.28		0.01						
66-67	0.22	0.25	0.05	0.02	0.04	0.01			0.03	
67-68	0.21	0.24	0.32	0.04	0.16	0.03			0.15	
68-69	0.16	0.21	0.81	0.11	0.28	0.26	0.01		0.34	
69-70	0.20	0.26	0.29	0.33	0.34	0.25	0.07		0.28	
70-71	0.23	0.26	0.17	0.33	0.37	0.34	0.14		0.19	
71-72	0.15	0.24	0.02	0.17	0.28	0.29	0.20		0.09	
72-73	0.20	0.25	0.15	0.03	0.19	0.30	0.18	0.02	0.21	
73-74	0.16	0.26	0.30	0.61	0.22	0.29	0.32	0.07	0.27	
74-75	0.18	0.20	0.36	0.32	0.34	0.30	0.23	0.18	0.26	
75-76	0.19	0.23	0.28	0.25	0.31	0.32	0.24	0.11	0.12	
76-77	0.17	0.24	0.32	0.19	0.30	0.30	0.23	0.29	0.12	
77-78	0.20	0.10	0.38	0.31	0.15	0.29	0.30	0.18	0.20	
78-79	0.15	0.15	0.27	0.25	0.17	0.26	0.18	0.25	0.28	
79-80	0.11	0.21	0.29	0.30	0.08	0.21	0.30	0.48	0.29	
80-81	0.01	0.09	0.36	0.13	0.00	0.35	0.28	0.70	0.21	
81-82	0.01	0.00	0.25	0.27	0.05	0.18	0.25	0.57	0.15	
82-83	0.13	0.06	0.29	0.35	0.22	0.22	0.25	0.63	0.22*	
83-84	0.15	0.02	0.28	0.28	0.25	0.25	0.28	0.35	0.18*	
84-85	0.17	0.07	0.29	0.26	0.23	0.21	0.23	0.63	0.22*	
Total discharge method B	3.30	3.71	5.51	4.58	3.99	4.66	3.69	4.46	3.82*	37.72
Total discharge method A	3.45	3.75	5.47	4.39	3.95	4.57	4.13	5.03	3.96*	38.69
Total discharge method C	3.09	3.69	5.55	4.66	3.80	4.50	3.70	4.82	3.67*	37.49

* Includes 1.25 × (discharge '82) to bring Hunterston to 31-3-83

† Fiscal years for Hunterston

$$P_c = \int_0^{B_{\max}} G(B) dD = \frac{M}{B_{\max}} \int_0^{B_{\max}} G(B) dB = P_d$$

where P_d is the plutonium discharged in the rise to the steady state. Plutonium discharged in steady state $P_s = D_s G(B_{\max})$.

Therefore total plutonium production (including plutonium in core) after substitution from (3) is given by

$$P_T = \frac{E_T G(B_{\max})}{B_{\max}} + M \left\{ \frac{2}{B_{\max}} \int_0^{B_{\max}} G(B) dB - G(B_{\max}) \right\} \quad (4)$$

Hence, if the ideal refuelling line was followed, the total plutonium production when total thermal energy E_T has been generated is determined in terms of one parameter, the steady-state burn-up $B_{\max}$.

In Table 3 we present the total plutonium discharged for each station by 31 March 1985 using a $B_{\max}$ which is the average of the discharge burnups in Method A for each station for the 10 years prior to 31 March 1985. When $B_{\max}$ is calculated for shorter periods or from equation (2) the totals for individual reactors vary in the range $\pm (0-3)$ per cent. This suggests that the method chosen to determine $B_{\max}$ is not

Table 4 Comparison of Method B predictions for plutonium production 1977-83 with CEGB figures for dispatched fuel 1978-84

	Plutonium produced during 1977-83 Method B (te)	Plutonium dispatched during 1978-84 CEGB (te)
Bradwell	0.605	0.594
Berkeley	0.613	0.637
Hinkley Point A	1.846	1.908
Trawsfynydd	1.625	1.565
Dungeness A	0.672	0.834
Sizewell	1.510	1.600
Oldbury	1.563	1.607
Total	8.434	8.745

critical.

An interesting result of Method C is that from equation (4) P_T consists of two parts: a first (larger) term which is proportional to E_T and a second (smaller) term which is always positive. The method used by Hesketh³⁴ and by Simpson¹ to estimate P_T assumes a quoted value for plutonium production per unit of electrical energy generated which for constant thermal efficiency means they were assuming plutonium production proportional to E_T . Since the second term in (4) is always positive, such estimates must be underestimates of P_T as Hesketh claimed. Our calculations suggest that the second term is approximately 2.4 te, summed over all the Magnox reactors.

The possible end uses of plutonium depend critically on its isotopic composition. We have calculated using Method B the plutonium discharged in two plutonium 240 purity bands 0–7 per cent and 0–15 per cent and the results are shown in Table 5. The 15 per cent figure is important because we know there is currently no plutonium of Pu 240 content less than 15 per cent in the civil stockpile³⁵. The plutonium of Pu 240 content less than 7 per cent would be particularly useful for military purposes, though plutonium of considerably worse purity could be blended with very high-purity plutonium to form acceptable weapons-grade plutonium. To put the numbers in Table 5 in perspective, Lovins states that the critical mass for weapons-grade plutonium with a reflector is less than 5 kg³⁶.

Plutonium balance

Plutonium must be 'lost' because reprocessing is not 100 per cent efficient. It may either be contained in solid or liquid waste, or discharged into the Irish Sea. By 1974 solid waste accumulated at Sellafield contained a little under half a tonne of plutonium¹³. It was anticipated that the corresponding figure for plutonium losses would not be so great over subsequent years. The Department of Energy have refused to answer parliamentary questions requesting an update of this figure³⁷ though they have admitted³⁸ that only about half of the quantity arises from CEBG and SSEB spent fuel.

From radiological data^{39,40} on discharges

Table 5 Production of high purity plutonium (te) for all CEBG and SSEB stations

	Plutonium 240	
	Sum all plutonium isotopes	
	0% – 15%	0% – 7%
Up to 31–3–69	2.3 ± 0.4	0.2 ± 0.1
From 31–3–69 to 31–3–71	0.8 ± 0.2	0.07 ± 0.02
After 31–3–71	1.1 ± 0.2	0.09 ± 0.05

of plutonium isotopes into the sea we calculate that approximately 280 kg of plutonium has been lost in this way up to end 1983. According to a recent parliamentary answer approximately 70 per cent of this arises from CEBG and SSEB spent fuel⁴⁰. In total we assume 0.5 ± 0.2 te of plutonium from the civil stockpile has been lost during reprocessing.

In Table 6 we show the total of plutonium in core and discharged according to Methods A, B and C at the four dates for which official information on plutonium stocks is available^{16,33,35,41}. We note that the three methods give differences of 2 per cent or less for the plutonium total at each date. This agreement suggests that uncertainties in exact refuelling policy including discharge figures for the 1960s are not very important when considering the total of plutonium produced. This is supported by the small changes in the total of -0.5 per cent when SUB figures rather than BEST (Table 2) are used in Method B, and $+1.7$ per cent when a uniform discharge throughout the year is assumed rather than mid-year discharge.

The three methods also calculate discharge burn-up differently. Hence the similarity of results suggests that systematic errors due to the lack of detailed knowledge of the burn-up variation within the reactor are probably smaller than the differences between the totals of the three methods.

All three methods assume the same $G(B)$ which for the civil reactors cannot be directly checked against published data. As a test of the sensitivity of our results to our choice of $G(B)$ we investigate the effect of using a worst-case $G(B)$ specified by parameters at the extent of the range which is reasonable: a fast fission factor of

1.033; $r = 0.07$; higher T_m where there is ambiguity in the literature. This $G(B)$ for a Calder Hall temperature lies well below the Tyrer curve and the agreement with the isotopic ratios in Table 1 worsens, but the plutonium total according to Method B falls by only 1.5 per cent. Given that our channel-averaged $G(B)$ underestimates the reactor averaged situation by approximately 1 per cent, we feel that the error on the missing plutonium given below accommodates such systematic effects.

According to Method B the total plutonium increase between 31 December 1981 and 31 March 1985 is 7.7 te. This agrees with the difference calculated from the parliamentary answers of 7.5 ± 0.5 te. Since an interpolation is required to produce plutonium totals at 31-12-81 it is probably safer to compare the difference between totals in 31-3-85 and 31-3-83 which is 4.9 te predicted by Method B and 5.0 ± 0.5 te in the parliamentary answers.

Missing plutonium

We conclude from Table 6 that the amount of plutonium unaccounted for is 6.8 ± 0.8 te according to our preferred Method B. After subtraction of the 0.5 ± 0.2 te lost in reprocessing (which is not included in the subtotal of civil stocks⁴²), the missing balance is 6.3 ± 0.8 te.

It is interesting to note that our estimate for the balance agrees with the figure of 6.667 te which was expected to be the maximum involved in the exchange¹ between the United Kingdom and the United States, based on costs in the US enabling act.

Previous studies of UK plutonium production have arrived at the following estimates for the balance of civil plutonium: Durie and Edwards⁴³, 14.5 te; Hesketh³⁴, 3.4 te; Simpson¹, 3.3 te. As discussed earlier we believe that approximately 2.4 te should be added to both the Hesketh and Simpson estimates. Their estimates would then be in agreement with ours.

Simpson favoured 3–4 te for the amount consigned to the US on the basis of the amount of plutonium produced by 1969. On 1–4–69 plutonium in their spent fuel arriving at Windscale became the property of the CEBG rather than the UK Atomic Energy Authority. The CEBG has stated^{44,45} that fuel reprocessed prior to 1

Table 6 Total plutonium (discharged and in core) from Methods A, B, & C and comparison with Parliamentary Answers

DATE	COL 2*	Method A		Method B		Method C	
	Sub-total of civil stockpile (te)	Total including core as in COL 2 (te)	Difference from COL 2 (te)	Total discharge +core (te)	Difference From COL 2 (te)	Total discharge +core (te)	Difference from COL 2 (te)
31-12-81	33.0	40.00	7.00	39.60	6.60	39.09	6.09
31-3-83	35.5	43.32	7.82	42.44	6.94	41.76	6.26
31-3-84	38.0	45.40	7.40	44.87	6.87	44.03	6.03
31-3-85	40.5	48.19	7.69	47.32	6.82	46.33	5.83
Mean difference			7.48		6.81		6.05

* Figures as quoted in Parliamentary Answers

April 1969 provided the plutonium sent to the United States. Our Method B gives 3.5 ± 0.4 te as the total produced by mid fiscal-year 1968–69, which is consistent with Simpson's estimate. If the plutonium unaccounted for is to be reduced to the amount produced by mid-year 1968–69 our calculations would need to overestimate production by 9 per cent. However, comparison with the CEGB dispatch data suggests that our calculations are a 3.6 per cent underestimate. Furthermore our figure cannot be a 9 per cent overestimate since the adjusted Method B total of plutonium in core on 31 March 1985 would then be 8.8 te which is incompatible with the figure of 9.5 te quoted in the parliamentary answer to the nearest half-tonne.

US plutonium use

According to the government⁴⁶ the bulk of the civil plutonium sent to the US is in the inventory of one fast research reactor, the zero power plutonium reactor (ZPPR), "in the core" of another, the fast flux test facility (FFTF), and "a sizeable quantity was used to make californium for medical purposes. The remaining small quantity is in use for experimental purposes elsewhere in the civil programme, for example at Argonne and Batelle."

ZPPR has an inventory of 3.8 te, of which a "portion" of the 3.4 te of fuel-grade plutonium came from the United Kingdom⁴⁷. FFTF only went into operation in 1981, has a core loading of 550 kg plutonium-239⁴⁸ and only a "small portion" of FFTF fuel was supplied by the United Kingdom⁴⁷. It has been estimated that at most a few hundred kilograms of its inventory of 2.9 te came from the United Kingdom⁴⁸. The amount of plutonium used for californium production has subsequently been revealed as 200 kg². If 200 kg is a "sizeable quantity" then the "remaining small quantity" in use at Argonne

and Batelle is insignificant.

We therefore estimate that the total UK plutonium in the destinations listed by the government is likely to be less than 4.0 te. The UK civil plutonium in these destinations could be considerably less if UK military plutonium was involved, as is possible⁴⁹.

Hence on the basis of our best Method B estimate we believe that at least 2.3 ± 0.8 te of UK civil plutonium is in destinations other than those given by the government.

Conclusions

The agreement between our three methods suggests that, despite the absence of public data on fuel discharges in the 1960s, it is possible to calculate the total plutonium produced by the civil Magnox reactors to a reasonable accuracy. We conclude that 6.3 ± 0.8 te of civil plutonium, approximately one-sixth of the total civil stockpile, are currently missing. We believe there is at least 2 te of UK civil plutonium in destinations other than those admitted in parliamentary answers. Until this is clarified the suspicion will exist that these destinations could be military.

Our calculations agree with the rather limited data available on plutonium production in civil Magnox reactors: parliamentary answers; isotopic ratios in CEGB dispatch data; and the "Flowers Report". Indeed they underestimate plutonium production when a comparison is made with CEGB dispatch figures.

In view of our findings we believe it is important that the UK government provides a much fuller explanation of the fate of civil plutonium produced during the 1960s, publishes more detailed information on civil plutonium production since 1971 and accepts effective safeguards on all civil nuclear facilities. This should include the currently unsafeguarded Mag-

nox reprocessing line at Sellafield which handles both civil and military plutonium and which has been the subject of continuing conflict between the government and EURATOM (the appropriate safeguards agency) since the United Kingdom joined the EEC⁵⁰.

We also find it most unsatisfactory that the government refuses to publish information on plutonium production by individual civil reactors even in recent years⁵¹, that this information is not supplied to EURATOM⁵², and that the CEGB removes the necessary data from its computer records⁵³.

Only by clarifying the extent of past links between civil and military nuclear programmes in the United Kingdom and by implementing procedures to prevent any such future re-occurrence can the government and the nuclear industry hope to strengthen the international non-proliferation regime. Such clarification would now be timely with the Non-Proliferation Treaty review conference under way in Geneva.

We are grateful to the Joseph Rowntree Charitable Trust and the Union of Concerned Scientists for providing financial assistance, to Scientists Against Nuclear Arms and the European Proliferation Information Centre for their support, to the London New Technology Network for the use of computing facilities, to Rob Edwards, Dafydd Elis Thomas M.P. and Tim Williams for seeking out information and to Zoe Saunders for help with computing. Useful discussions were also held with Norman Dombey, Professor P.J. Grant, Ross Hesketh and John Simpson. □

K. W. J. Barnham is at the Blackett Laboratory, Imperial College, London SW7 2BZ, UK; D. Hart is in the School of Environmental Sciences, University of East Anglia, Norwich NR4 7TJ, UK; J. Nelson is at London New Technology Network, London NW1 9ES, UK; R. A. Stevens is at Intel Corporation, Santa Clara, California 05051, USA.

1. Simpson, J. *The Independent Nuclear State*, (Macmillan, London, 1983); Dombey, N. *The Guardian*, London, 3-12-81.
2. Official Report of the House of Commons, Hansard (hereinafter referred to as OR), 9-3-83, Col 402.
3. Sizewell Inquiry Documents, CND/S/151, CND/S/167 (ADD1, ADD2, ADD3), (1984).
4. Sizewell Inquiry Documents, CEGB/P/1 (ADD12, ADD13), (1984).
5. Grant, P. J. *Elementary Reactor Physics*, (Pergamon, Oxford 1966); Jakeman, D. *Physics of Nuclear Reactors*, E.U.P., London, (1966).
6. James, M.F. *J. Nucl. Energy* **23**, 517-536, (1969).
7. AEEW-R 94, UKAEA Winfrith, (1961).
8. *Directory of Nuclear Reactors*. IAEA, Vienna: Vol IV, (1962); Vol VII, (1968); Vol X, (1976).
9. Tyrer, J.G. *Developments in the Physics of Nuclear Power Reactors*, 47-93, IAEA, (1973).
10. Griggs, C.F. & Harper, R.G., *TRG Report*, 958(W), UKAEA Risley, (1965).
11. *Nuclear Engineering International*, Power Reactor Supplement (Annually).
12. Duckworth, F.C. & Greenwood, D.R. *Nuclear Energy* **20**, 133-139, (1981).
13. Royal Commission on Environmental Pollution, 6th Report, Cmnd 6618, (1976).
14. Sizewell Inquiry Documents, CND/S/167, (1984).
15. O.R., 12-5-83, Col 377.

16. O.R., 25-1-85, Col 545.
17. O.R., 4-7-85, Col 198.
18. Communications from Miller, D. R., Chairman, SSEB to Edwards, R. 24-10-83 and 22-3-84.
19. O.R., 22-2-85, Col 617.
20. O.R., 18-7-85, Col 247.
21. Evans, D.T. et al. *J. Br. Nucl. Energy Soc.* **4**, 217-226, (1965).
22. Moore, R.V. *J. Br. Nucl. Energy Soc.*, **8**, 123-132, (1969).
23. *Nuclear Engineering International* **14**, 737-739 (1969).
24. Harris, J.E. et al. *Proc. 4th U.N. Conf., Geneva 1971* **10**, 167-178, (1971).
25. Dixon, F. & Simmons, H.K. *J. Br. Nucl. Energy Soc.* **13**, 9-38, (1974).
26. *Proc. Instn. Mech. Engrs.* **183**, Pt 3G, (1969).
27. Communications from Drew, R.K. CEGB to Edwards, R., 21-11-83 and 30-3-84.
28. O.R., 5-7-85, Col 308.
29. 1970-74, Ref. [30]; 1979-80, Ref. [31]; 1981-85, Ref. [11]; Berkeley and Bradwell, 1968-69, Ref. [32]; Hunterston, 1965-68, Ref. [26].
30. *Operating Experience with Nuclear Power Stations in Member States*, IAEA, Vienna, (Annually).
31. Bindon, F.J.L., *Nucl. Energy*, **20**, 127-132, (1981).
32. Harris, J.E. *CEGB Research*, Dec 1974, 29-35, (1974).
33. O.R., 23-7-85, Col 473.
34. Hesketh, R.V., *Science and Public Policy*, **9**, 64-70, (1982).
35. O.R., 27-7-83, Cols 439-440.

36. Lovins, A.B., *Nature*, **283**, 817-823, (1980).
37. O.R., 20-3-85, Col 516.
38. O.R., 15-3-85, Col 301.
39. *The Risks of Leukaemia and other Cancers in Seascale from Radiation Exposure*, NRPB-R171, National Radiation Protection Board (1984).
40. O.R., 4-7-85, Col 219.
41. O.R., 1-4-82, Col 168.
42. O.R., 20-3-85, Col 516.
43. Durie, S. Edwards, R. *Fuelling the Nuclear Arms Race* Pluto, London, (1982).
44. Communication from Bonner, F.E. CEGB to Hesketh R.V., 26-5-82.
45. Sizewell Inquiry Transcript, Day 283, p.72G (1984).
46. O.R., 27-7-82, Cols 438-9.
47. Communication from Hodel D. P. US Secretary of Energy to Ottinger, R. L. 5-3-84.
48. Communication from Hoenig, M. to Edwards, R., 3-5-84, and Sizewell Inquiry Documents CND/P/1 (1984).
49. Avery, D.G., Sizewell Inquiry Transcript, Day 274, p.53B, (1984).
50. Sizewell Inquiry Transcript, Day 274, p.61B-65G (1984).
51. Communication from Walker, P. Secretary of State for Energy to Dafydd Elis Thomas MP, 11-6-85 and Ref [16].
52. O.R., 1-4-85, Col 465.
53. Sizewell Inquiry Transcript, Day 193, p.14H (1984).

The galactic ridge observed by Exosat

R. S. Warwick, M. J. L. Turner, M. G. Watson & R. Willingale

X-ray Astronomy Group, Physics Department, University of Leicester, Leicester LE1 7RH, UK

We have used Exosat to map the distribution of X-ray emission in the galactic plane. In addition to numerous point sources, the map reveals a narrow continuous ridge of emission which extends along the galactic plane either side of the galactic centre out to $\sim 40^\circ$ longitude. The origin of this feature remains uncertain although a major contribution from luminous discrete X-ray sources can be excluded.

REPORTS of the discovery of extended hard X-ray emission from the Galaxy date back to early observations with sounding rockets^{1,2}. Although many of these first measurements were contaminated by bright discrete sources, evidence for an excess X-ray flux from the galactic plane between 50° and 65° longitude, which is not explained in terms of known discrete sources, has remained convincing^{3,4}. More recently, the comprehensive study of the X-ray background by the A2 experiment on HEAO-1 (ref. 5) has revealed the existence of two galactic components. One is a large scale height feature first observed in Ariel V⁶ and Uhuru^{7,8} data, and is evident at high galactic latitudes ($|b| > 10^\circ$). This component seems to have a scale height > 1 kpc, a thermal spectrum with $kT \sim 9$ keV and an X-ray luminosity of $\sim 3 \times 10^{38}$ erg s⁻¹ (ref. 9). The HEAO-1 A2 instrument also detected a narrow ridge of X-ray emission coincident with the galactic plane¹⁰ that Worrall and Marshall¹¹ suggest arises from a low-luminosity ($L_x < 4 \times 10^{32}$ erg s⁻¹) source population comprising RS CVn, cataclysmic variables and possibly dM stars. But the HEAO-1 measurements excluded the entire region of the galactic plane within 50° of the galactic centre because of source confusion in the $3 \times 1.5^\circ$ beam (full width at half maximum, FWHM) of the A2 instrument. Here we report the detection of a much brighter galactic ridge component within $\pm 40^\circ$ of the galactic centre.

Exosat galactic plane survey

The Exosat medium-energy (ME) proportional counter array has a sensitive area of $\sim 1,600$ cm² and a collimated field of view of $0.75 \times 0.75^\circ$ (FWHM)¹². In the normal mode of operation the spacecraft attitude and orbit control system provides a pointing accuracy of ~ 5 arc s but is also capable of performing controlled slews with attitude reconstruction to ~ 1 arc min from gyro data. In addition, the highly eccentric orbit allows uninterrupted observations for 72 h out of each 90 h orbit. Thus, the successful launch of Exosat provided a unique opportunity to map the distribution of galactic X-ray emission above 2 keV with good sensitivity and positional accuracy and a relatively narrow beam. The survey was conducted with half of the ME detectors in a 45 arc min offset pointing configuration (to improve sky coverage) and was performed in segments during the period

14–28 June 1984 so as to coincide with the passage of the anti-sun through the galactic centre region. Each survey segment consisted of a 90° slew at a rate of 42° h⁻¹ along a great circle from a starting point on the galactic plane at galactic longitude (l) either $+96^\circ$ or 276° up to the anti-sun position, followed by a short (0.75°) slew in an orthogonal direction and finally a return slew leg to the starting position. The full survey consisted of 15 such segments and resulted in the mapping of a lens-shaped region 10° wide at the galactic centre decreasing to 2° wide at $l = 96^\circ$ or 276° . An intensity map was produced from the raw data by accumulating counts in $20' \times 20'$ pixels over the whole area mapped and then dividing by the appropriate exposure time. An average background count rate was obtained from the data recorded in regions excluding obvious discrete sources and subtracted from the map. Figure 1 shows the central part of the map obtained in the 2–6 keV energy band.

Galactic ridge

Apart from many point sources, the most striking feature of this map is the band or ridge of X-ray emission of width $\sim 2^\circ$ (FWHM), which extends along the galactic plane on both sides of the galactic centre to $|l| \approx 40^\circ$; the map also shows extended emission in the region of the galactic bulge. (A catalogue of the point sources detected in the survey will be published elsewhere.) Figure 2a shows the distribution in longitude of the X-ray emission from a narrow cut at $b = 0^\circ$. The underlying distribution in the region $|l| \leq 40^\circ$, on which many bright discrete sources appear, corresponds to the unresolved emission which constitutes the galactic ridge. The ridge is least contaminated by obvious discrete sources in the longitude range $l = 20$ – 40° (Fig. 2b), but even in this region there are significant fluctuations with respect to a smooth intensity distribution. Some of the bright spots in this part of the ridge are almost certainly individual discrete sources that do not appear in existing medium-energy X-ray catalogues. For example, the intensity enhancement at the position of the supernova remnant W44 (G34.6–0.5) implies a 2–6 keV flux for this object which is consistent with Einstein measurements¹³. The bulk of the unresolved emission which constitutes the galactic ridge feature apparent in Fig. 1 cannot, however, be caused by a few sources,

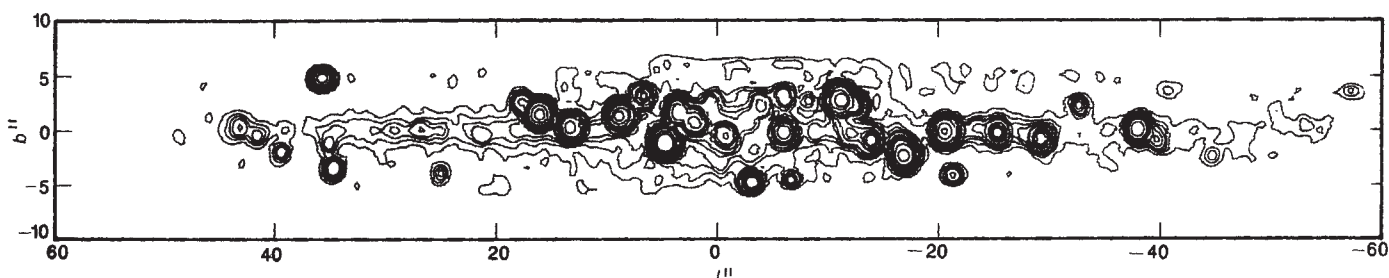


Fig. 1 Contour map of the 2–6 keV X-ray emission from the galactic plane in the region $|l| < 60^\circ$. The lowest contour level corresponds to an X-ray intensity of $\sim 7 \times 10^{-12}$ erg s⁻¹ cm⁻² per beamwidth. The contour levels in these units are 1, 2, 3, 4, 6, 8, 16, 24, 32, 40, 80, 200 and 400. The map has been smoothed by cross-correlation with a function which approximates the ME collimator response.

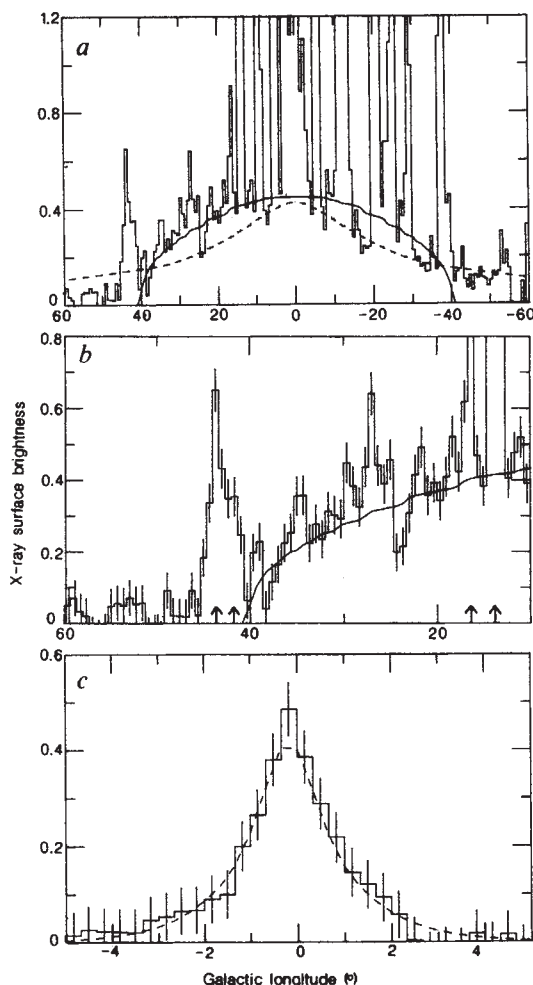


Fig. 2 *a*, X-ray intensity as a function of galactic longitude in a narrow band of width 1.3° (4 pixels) at $b = 0^\circ$ for the region $|l| < 60^\circ$. The predicted profiles for the finite disk (solid line) and exponential disk (broken line) models of Table 1 are indicated. *b*, Longitude profile for the region $l = 10$ – 60° . Solid curve, finite disk model of Table 1. Emission from bright catalogued sources is indicated by the arrows. *c*, Average latitude profile of the galactic ridge in the longitude range $l = 20$ – 30° . Dashed line, finite disk model of Table 1. Units of intensity are $7 \times 10^{-11} \text{ erg s}^{-1} \text{ cm}^{-2}$ per beamwidth (2–6 keV).

given the close coincidence of the ridge with the galactic plane and its apparent extent over $\sim 80^\circ$ of longitude. Moreover, profiles in galactic latitude (Fig. 2c) indicate that the ridge is consistently wider than the effective instrument response.

Spectral data can, in principle, place important constraints on any model for the origin of the ridge emission. Unfortunately, the present Exosat measurements have a rather poor statistical precision because of the very short effective exposure time. In terms of a thermal bremsstrahlung model a temperature of 6_{-2}^{+14} keV is indicated (photon index $\Gamma = 2.2_{-1.1}^{+2.4}$ in a power law model) with a low energy cut-off consistent with the models described below.

Modelling X-ray emission

We have modelled the observed X-ray emission in terms of a galactic component with volume emissivity $\eta(R, z)$ which is a function only of the galactocentric radius R and height above the galactic plane z . We consider two models: (1) a finite disk of radius R_D within which $\eta = \eta_D \exp(-|z|/z_s)$; (2) an exponential disk where $\eta = \eta_0 \exp(-R/R_s) \exp(-|z|/z_s)$. For the former case the rapid fall-off in the longitude distribution for $|l| \geq 40^\circ$ implies $R_D \sim R_0 \sin 40^\circ \approx 6.5 \text{ kpc}$, where R_0 is the distance of the galactic centre (taken to be 10 kpc); the width of the ridge

($\sim 2^\circ$) then requires $z_s \sim 100 \text{ pc}$. Photoelectric absorption in the 2–6 keV band is not negligible for large distances through the galactic disk, and we have included in the modelling the effects of a gas distribution which approximates to the known distribution of both neutral and molecular hydrogen (assuming cosmic element abundances). For the H I we assume an extended disk with an exponential scale height of 120 pc and a constant density on the galactic plane of 0.5 atom cm^{-3} (refs 14, 15). For the H_2 we use the densities and radial distribution derived by Sanders *et al.* from CO measurements¹⁵, and assume a half-width for the molecular layer of 70 pc. Although the total mass of H_2 in the inner Galaxy remains controversial^{16,17}, the CO measurements presumably provide a rather more direct measure of the atomic densities relevant to X-ray absorption. The optical depth in the 2–6-keV band to the galactic centre with this absorption model is ~ 1 .

Table 1 gives the parameters for both models which have been used to derive the longitude and latitude profiles shown in Fig. 2. We have not attempted to model the obvious asymmetry in the observed longitude distribution or any other fine detail. Neither have we attempted to constrain the model parameters by minimizing χ^2 . We simply use the models to illustrate the basic form of the observed ridge emission. Note that if absorption related to the H_2 component is excluded, the values of η_D and η_0 given in Table 1 decrease by a factor of ~ 1.9 and ~ 2.2 , respectively, while the other parameters of the model are little affected.

The predicted emission profiles for the finite disk model (Fig. 2) are in reasonable agreement with the observations. In several directions there is an apparent deficit of observed emission compared with the assumed model which may be indicative of large-scale spatial variations of the X-ray emissivity of the ridge or alternatively of substantial variations in the line-of-sight absorption. CO measurements suggest that variations in the X-ray absorption in the galactic plane will give rise to a significant modulation of the observed X-ray intensity and in fact there is evidence for an anti-correlation of CO emission line strength with X-ray intensity; for example, the dip at $\sim 24^\circ$ in Fig. 2b corresponds to a local maximum in the CO emission¹⁵. The exponential model, which was chosen to be more representative of the matter distribution in the galactic disk (for example, ref. 18), provides only a poor fit to the observed longitude distribution (Fig. 2a). With this model any reasonable choice of normalization to fit the observed X-ray emission from the inner Galaxy results in a predicted intensity for $|l| > 40^\circ$ which is well in excess of that observed.

Note that the galactic ridge emission detected by HEAO-1 A2 at $|l| > 50^\circ$ has a volume emissivity an order of magnitude smaller than that derived here for the $|l| < 40^\circ$ region and also a larger scale height (corresponding to a disk of half-width $244 \pm 22 \text{ pc}$; ref. 10). The HEAO-1 A2 component is below the detection threshold of the Exosat observations. Furthermore, it is uncertain whether the two features are related.

Discrete source origin

We first consider the possibility that the galactic ridge observed by Exosat originates in a population of unresolved discrete sources of luminosity $L_x \text{ erg s}^{-1}$ (2–6 keV), and number density n where $n = \eta/L_x \text{ pc}^{-3}$. The number of sources N contributing

Table 1 Parameters used in modelling the galactic ridge emission

Finite disk	Exponential disk
$\eta = \eta_D \exp(- z /z_s) \quad R < R_D$	$\eta = \eta_0 \exp(-R/R_s) \exp(- z /z_s)$
$\eta = 0 \quad R > R_D$	
$R_D = 6.5 \text{ kpc}$	$R_s = 3.5 \text{ kpc}$
$z_s = 100 \text{ pc}$	$z_s = 100 \text{ pc}$
$\eta_D = 6.2 \times 10^{27} \text{ erg s}^{-1} \text{ pc}^{-3}$	$\eta_0 = 1.1 \times 10^{28} \text{ erg s}^{-1} \text{ pc}^{-3}$
$L_x (\text{Galaxy}) = 1.5 \times 10^{38} \text{ erg s}^{-1}$	$L_x (\text{Galaxy}) = 1.7 \times 10^{38} \text{ erg s}^{-1}$

to the ridge per beam width in longitude (taking $l = 25^\circ$ as a representative direction) in the finite disk model is then

$$N = 7.5 \times 10^{35} L_x^{-1}$$

If we take $N > 3$, this requires $L_x < 2.5 \times 10^{35} \text{ erg s}^{-1}$. In principle, the fluctuations about a smooth distribution in longitude (at least for the region $l = 20\text{--}40^\circ$; Fig. 2b) suggest $N < 30$, which would imply $L_x > 2.5 \times 10^{34} \text{ erg s}^{-1}$. This argument will not apply, however, if the observed fluctuations are largely the result of only a very few sources of flux $\sim 5 \times 10^{34} \text{ erg s}^{-1} \text{ cm}^{-2}$ which are coincident with the ridge but not representative of the population which accounts for its $\sim 80^\circ \times 2^\circ$ extent.

A more stringent constraint comes from a comparison with Einstein observations in the galactic plane. We have calculated (Fig. 3) the number of point-like discrete sources N' per square degree (again taking $l = 25^\circ$, $b = 0^\circ$ as a representative direction) which would be detectable in the softer band of the Imaging Proportional Counter (IPC) in a typical Einstein observation (IPC count rate of $\geq 0.01 \text{ c.p.s.}$). Our analysis of 14 Einstein pointings in the region of $l = 24\text{--}26^\circ$, $b = \pm 1^\circ$, which are included in the galactic plane survey of Hertz and Grindlay¹⁹, indicates $N' < 0.5$. This low detection rate by the IPC implies that a source population with luminosity L_x in the range $10^{33.5} \text{ erg s}^{-1} \leq L_x \leq 10^{36} \text{ erg s}^{-1}$ can be excluded as a major contributor to the galactic ridge.

In summary, analysis shows that although a few relatively bright but unresolved sources probably contribute to the ridge and give rise to significant fluctuations in its intensity, the bulk of the observed emission must arise, assuming a discrete source origin, in a relatively low-luminosity X-ray source population (that is, $L_x < 10^{33.5} \text{ erg s}^{-1}$). Furthermore, the derived scale height of $\sim 100 \text{ pc}$ firmly indicates an association with population I, although the radial distribution in the Galaxy does not closely correspond to that of either spiral arm or smooth disk population material. The luminosity restriction implies that at least 3×10^4 sources must be contributing to the ridge. This immediately excludes, for example, young supernova remnants (age $< 5,000 \text{ yr}$) on the assumption that the galactic supernova rate is $\sim$ one per 30 yr ^{20,21}. In any case, supernova remnants have thermal spectra which are generally too soft to account for the observed ridge spectrum. Classes of X-ray sources which have roughly correct luminosity and spectral characteristics include cataclysmic variables, RS Cvn systems and possibly Be stars such as X Per and $\gamma \text{ Cas}$ ¹¹.

Diffuse origin

An alternative possibility to a discrete source origin is that the bulk of the ridge emission arises in a diffuse process. The X-ray volume emissivity due to inverse Compton scattering of the microwave background, far infrared photons and starlight, derived from estimates of the interstellar cosmic-ray electron spectrum (for example, ref. 22), falls short of the requirement by at least two orders of magnitude^{6,10}. In addition, the scale height of $\geq 1 \text{ kpc}$ for cosmic-ray electrons obtained by modelling radio^{23,24} and γ -ray measurements²⁵ does not match that of the X-ray emission. Medium-energy X-rays are also produced as synchrotron radiation from $\sim 10^{14} \text{ eV}$ electrons moving in the galactic magnetic field. In a lifetime of $\sim 10^4$ years such particles travel $< 170 \text{ pc}$ by diffusion¹⁰ and might fill a disk in the inner Galaxy of the form observed. However, an optimistic extrapolation of the cosmic-ray electron spectrum, at least in the inner Galaxy region, is required for this process to be significant and the energy input needed to sustain the requisite electron population would be large¹⁰.

The question of whether the X-ray emission has a diffuse thermal origin arises following the observation²⁶ by Tenma of intense Fe line emission at 6.7 keV from several directions in the galactic plane in the longitude range $280\text{--}340^\circ$. A thermal bremsstrahlung continuum was also detected, with a temperature ranging from 5 to 9 keV depending on position. Given the relatively low angular resolution of the Tenma observations (collimator FWHM $\sim 3^\circ$), the possibility that this line and ther-

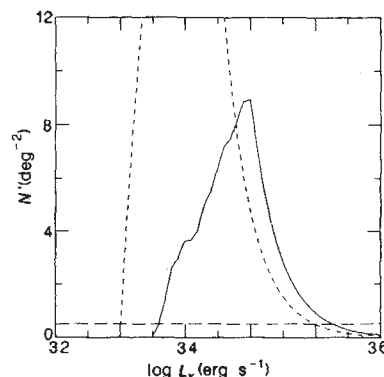


Fig. 3 Predicted number, N' , of sources per square degree detectable by the Einstein IPC above a limiting sensitivity of 0.01 IPC c.p.s. as a function of the source luminosity L_x at $2\text{--}6 \text{ keV}$. The finite disk model is assumed with (solid curve) and without (dashed curve) the H_2 component of absorption. The value of N' plotted has been averaged over a $1.5 \times 1.5^\circ$ region centred on $(l, b) = (25^\circ, 0^\circ)$. Broken line, observational limit, that is, $N' < 0.5$. The calculation assumes a power law source spectrum with photon index of 1.5 and no low-energy cut-off intrinsic to the source.

mal continuum originates in a narrow galactic ridge component (as observed by Exosat) cannot be excluded. The particle density required to produce the observed $2\text{--}6 \text{ keV}$ volume emissivity in the finite disk model (assuming $kT \sim 6 \text{ keV}$ and a filling factor ~ 1) is $n \sim 3 \times 10^{-3} \text{ cm}^{-3}$, with an associated pressure $P/k \sim 4 \times 10^5 \text{ cm}^{-3} \text{ K}$, which is substantially larger than that of the hot component of the interstellar medium²⁷. This plasma seems to be too hot to be either gravitationally bound or in pressure confinement and adiabatic expansion out of the disk would provide the dominant cooling mechanism (the radiative cooling timescale is $> 10^{10} \text{ yr}$). Presumably, a high-temperature phase of a supernova-shocked interstellar medium must be invoked. A direct test of this rather radical hypothesis is to search, using Exosat, for Fe line emission at positions along the galactic ridge.

Conclusions

In conclusion, the Exosat galactic plane observation has revealed a new feature of our Galaxy. The narrow ridge of emission, apparent in Fig. 1, has an X-ray luminosity of $\sim 10^{38} \text{ erg s}^{-1}$, a scale height of $\sim 100 \text{ pc}$ and a radial extent within the Galaxy of $\sim 6.5 \text{ kpc}$. This inferred spatial distribution is dissimilar to that of the bright X-ray sources in the Galaxy and does not closely resemble the distribution of known population I tracers. Whether this component originates in a low-luminosity source population ($L_x < 10^{33.5} \text{ erg s}^{-1}$), a diffuse mechanism or a combination of processes is at present unclear. If the bulk of the emission is caused by hot ($10^7\text{--}10^8 \text{ K}$) gas, there would be major consequences for our understanding of the dynamics and heating of the interstellar medium.

We thank our colleagues, particularly L. Jones, Professor K. A. Pounds and Dr G. Stewart for helpful comments and discussions. We also thank the Exosat observatory team for carrying out the complicated set of manoeuvres which comprised the galactic plane scan observation.

Received 20 May; accepted 31 July 1985.

- Cooke, B. A., Griffiths, R. E. & Pounds, K. A. *Nature* **224**, 134–137 (1969).
- Hudson, H. S., Peterson, L. E. & Schwartz, D. A. *Nature* **230**, 177–179 (1971).
- Bleach, R. D., Boldt, E. A., Holt, S. S., Schwartz, D. A. & Serlemitsos, P. J. *Astrophys. J. Lett.* **174**, L101–106 (1972).
- Wheaton, W. A. thesis, Univ. California, San Diego (1976).
- Rothschild, R. E. *et al. Space Sci. Instrum.* **4**, 269–301 (1979).
- Warwick, R. S., Pye, J. P. & Fabian, A. C. *Mon. Not. R. astr. Soc.* **190**, 243–260 (1980).
- Schwartz, D. A. in *X-ray Astronomy* (eds Baity, W. A. & Peterson, L. E.) 453–465 (Pergamon, Oxford, 1979).
- Protheroe, R. J., Wolfendale, A. W. & Wdowczyk, J. *Mon. Not. R. astr. Soc.* **192**, 445–454 (1980).
- Iwan, D. *et al. Astrophys. J.* **260**, 111–123 (1982).
- Worrall, D. M., Marshall, F. E., Boldt, E. A. & Swank, J. H. *Astrophys. J.* **255**, 111–121 (1982).
- Worrall, D. M. & Marshall, F. E. *Astrophys. J.* **267**, 691–697 (1983).
- Turner, M. J. L., Smith, A. & Zimmerman, H. U. *Space Sci. Rev.* **30**, 513–524 (1981).
- Smith, A. *et al. Mon. Not. R. astr. Soc.* (in the press).
- Baker, P. L. & Burton, W. B. *Astrophys. J.* **198**, 281–297 (1975).

15. Sanders, D. B., Solomon, P. M. & Scoville, N. Z. *Astrophys. J.* **276**, 182–203 (1984).
16. Lebrun, F. *et al. Astrophys. J.* **274**, 231–236 (1983).
17. Bhat, C. L., Issa, M. R., Houston, B. P., Mayer, C. J. & Wolfendale, A. W. *Nature* **314**, 511–515 (1985).
18. Bahcall, J. N. & Soneira, R. M. *Astrophys. J. Suppl.* **44**, 73–110 (1980).
19. Hertz, P. & Grindlay, J. E. *Astrophys. J.* **278**, 137–149 (1984).
20. Tammann, G. A. in *Supernovae: A Survey of Current Research* (eds Rees, M. J. & Stoneham, L. R. J.) 371 (Reidel, Dordrecht, 1982).
21. Lyne, A. G., Manchester, R. N. & Taylor, J. H. *Mon. Not. R. astr. Soc.* **213**, 613–639 (1985).
22. Webber, W. R., Simpson, G. A. & Cane, H. V. *Astrophys. J.* **236**, 448–459 (1980).
23. Baldwin, J. E. in *The Structure and Content of the Galaxy and Galactic Gamma Rays* (NASA CP-002, 1976).
24. Strong, A. W. *Mon. Not. R. astr. Soc.* **181**, 311–322 (1977).
25. Kniffen, D. A. & Fichtel, C. E. *Astrophys. J.* **250**, 389–397 (1981).
26. Koyama, K. in *X-ray Astronomy '84* (eds Oda, M. & Giacconi, R.) 325–328 (1984).
27. Shapiro, P. R. & Field, G. D. *Astrophys. J.* **205**, 762–765 (1976).

Continental sources of halocarbons and nitrous oxide

Michael J. Prather*

Center for Earth and Planetary Physics, Harvard University, Cambridge, Massachusetts 02138, USA

Estimates of continental sources of CFC-11, CFC-12, CCl₄, CH₃CCl₃ and N₂O are derived from the atmospheric lifetime experiment in Adrigole, Ireland, and anthropogenic emissions of CCl₄ and N₂O from Europe have been identified. Relative source strengths are consistent with global budgets for the halocarbons and N₂O. Different industrial release patterns for halocarbons are observed for Europe, the western United States and Australia.

RECORDS of the composition of the atmosphere over the past two decades^{1–9} show a global increase in the concentrations of CO₂, CH₄, N₂O, CO, chlorofluorocarbons-11 and -12 (CFC-11 and CFC-12) as well as many other halocarbons. Observations have often focused on clean-air, baseline conditions where a reliable trend in the global burden of these gases may be determined. Here, I analyse data for trace gases from a once polluted site (Adrigole, Ireland) to identify and quantify sources of N₂O and halocarbons from the European continent. Globally, these gases represent the major stratospheric supply of odd nitrogen and inorganic chlorine, which, in turn, control the abundance of ozone. Increasing levels of chlorine and odd nitrogen may lead to large reductions in stratospheric ozone^{10,11}.

Sources of atmospheric trace compounds such as N₂O and halocarbons are poorly quantified. Emission inventories for halocarbons are derived through surveys and estimates of global manufacture and distribution which contain major uncertainties^{9,12}. The budget for N₂O (ref. 7) is based on calculated losses from the stratosphere, on global scaling of few measurements of fluxes from soils and oceans, on the observed rate of increase in atmospheric concentration and on the detection of N₂O as a by-product of combustion^{13,14}. I derive here an accurate calibration of the relative emission of N₂O and halocarbons from Europe. This information is contained in the measurements of CFC-11 (CFCl₃), CFC-12 (CF₂Cl₂), CCl₄, CH₃CCl₃ and N₂O from the atmospheric lifetime experiment (ALE) station at Adrigole, Ireland⁹.

The ALE data have been presented in the form of monthly mean concentrations from which the pollution events have been removed⁹. The initial study was extensive and used these clean-air, baseline values to calculate highly accurate hemisphere and global trends to compare with estimates for industrial emissions, and thence to derive globally averaged atmospheric lifetimes. I examine here the variations of trace gas concentrations occurring on short timescales, from 1 to 10 days, which result from the passage of polluted air over the observing site. The complete set of individual observations from the first 3 yr at Adrigole has recently become available and is shown in Fig. 1. The ALE programme has noted that ~30% of these measurements at Adrigole seem to be polluted, containing significantly higher concentrations of specific halocarbons associated with anthropogenic sources in Europe. The ALE analysis has selectively removed these events; here, I focus on them, using the elevated concentrations during pollution events to derive relative source strengths of these gases from western Europe. Similar,

but less detailed, analyses have been reported for urban sources of methane¹⁵ and for CFCs in Australia (P. J. Fraser, personal communication). The site at Adrigole is particularly advantageous for sampling European land sources with respect to the relatively uniform oceanic background. ALE data from Cape Meares, Oregon and Cape Grim, Tasmania do not define continental emissions as clearly as the Adrigole data, but do indicate differing relative sources of these gases from western Europe, North America and Australia.

ALE measurements and analysis

The first 3 yr of measurements from the Adrigole station, spanning 1 July 1978 to 30 June 1981, are presented in Fig. 1. The gas chromatograph used two different columns, silicone and porosil, and values for CFC-11, designated S and P, were reported for both⁹. The data show numerous pollution episodes which seem to correlate across all gases except N₂O. Irregular intervals in the record complicate the time-series analysis.

In the current analysis, we have processed the ALE data for each of the 'six' gases individually as follows:

- (1) Original measurements, taken several times a day, are recorded to the nearest minute of observations, and have been temporally degraded into intervals of 8 h; where more than one observation falls within the 8-h interval, they are averaged.
- (2) A linear, least-squares regression was fitted to all of the 8-h data; all points were weighted equally and no attempt was made to fill gaps in the record.
- (3) These 'detrended' data were sorted, resulting in the probability distribution functions shown in Fig. 2.
- (4) A limit on extreme values was selected from the 0.5% and 99.5% probability levels; and those data in the upper/lower 0.005 fractile were rounded down/up to the chosen extrema.
- (5) The truncated data set was again fitted with a linear trend which was again removed.
- (6) Auto-correlations and cross-covariances of the residual variances were computed with a resolution of 8 h; again, no attempt was made to fill the numerous gaps in the data. Covariances (and auto-correlations) have been defined as

$$C_{12}(\Delta) = \frac{1}{N} \sum f_1(t) \times f_2(t + \Delta)$$

where there are N such pairs of residual trace gas mixing ratios, f_1 and f_2 , available at a time lag Δ .

(7) The coherence of a pair of gases (1, 2) is defined in the form of a normalized covariance,

$$h_{12}(\Delta) = c_{12}(\Delta) / [c_{11}(\Delta) \times c_{22}(\Delta)]^{1/2}$$

where the relative covariances (c_{12}) represent the increase in

* Present address: NASA/Goddard Institute for Space Studies, 2880 Broadway, New York, New York 10025, USA.

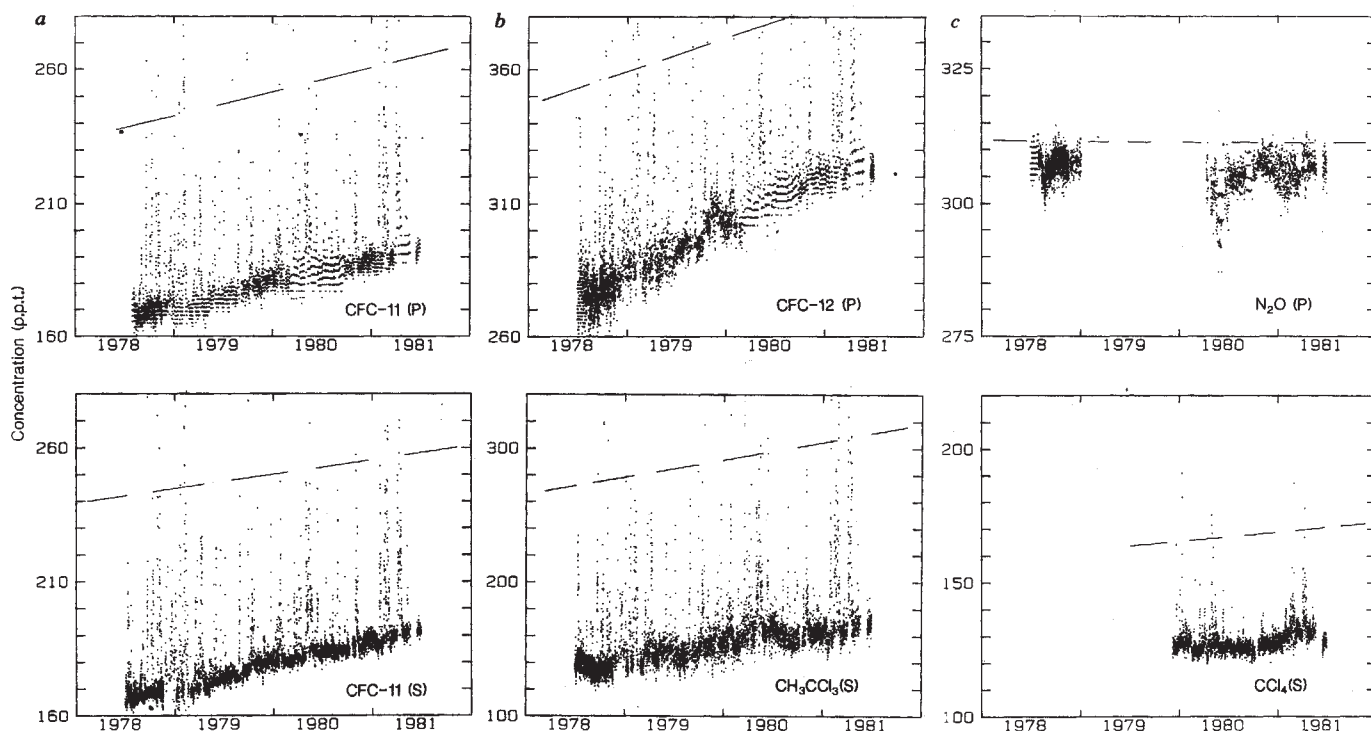


Fig. 1 ALE measurements from Adrigole, Ireland. All primary observations from the P-column (CFC-11, CFC-12, N_2O) and the S-column (CFC-11, CH_3CCl_3 , CCl_4) which fall within the ranges of the graphs have been plotted. Dashed lines indicate the 99-percentile level.

variance of short-term pollution events with respect to a baseline defined by the 20–30-day lag.

$$c_{12}(\Delta) = C_{12}(\Delta) - \langle C_{12} \rangle_{20-30\text{-day}}$$

This definition of coherence is meaningful only when C is significantly greater than the baseline, and thus values in Table 2 are reported only for lags out to 8 days where the pollution events are clearly distinguishable above the baseline concentrations of clean air.

Steps 4–5 were adopted because variances in the halocarbon data tend to be dominated by relatively few observations with very high values. This approach effectively filters the extreme, short-lived, infrequent pollution events into manageable variations in trace gas concentrations, without losing the basic information that the air was highly polluted. Repeating this analysis with limits placed at the 1.5 and 98.5% probability levels shows negligible sensitivity to these limits.

Trends in concentration for the trace gases as calculated in steps (2) and (5) are shown together with results from the original ALE analysis in Table 1. Except for CFC-11S these trends agree well and indicate that inclusion of the pollution events (removed by the ALE investigation in the original analysis) do not seem to distort the baseline trend observed at Adrigole. The calculated trend in CFC-11S is influenced by a few extremely high concentrations early in the record: truncating the distribution at $\pm 1.5\%$ probability raises the calculated trend to $+4.6 \pm 0.3\% \text{ yr}^{-1}$, closer to the other trends derived for CFC-11.

The cumulative probability distributions shown in Fig. 2 reveal two distinct categories of data. The linear region below the 60% level is indicative of a normal gaussian distribution that we assume represents the measurement of background air as influenced by instrumental variations plus short-term and seasonal variations not removed by the linear trend. Second, the steep rise in concentrations at the upper quartile of the probability distribution (for all gases but N_2O) points to sources of air with highly enriched concentrations, agreeing with the ALE observation that Adrigole experiences polluted European air $\sim 30\%$ of the time. These distributions may be used to derive relative source strengths for these gases as described below.

The temporal correlation of variations in the five ALE trace gases provide valuable information on their concurrency as well as on the duration of pollution events. The former is addressed here as we attempt to extract relative emissions integrated over a continent; the latter provides important tests for atmospheric models of tracer transport. Our focus is restricted to ± 30 -day lags because of interest in synoptic-scale variability. The original data exhibit discontinuities on a monthly or seasonal basis which make the interpretation of longer correlations more difficult (see

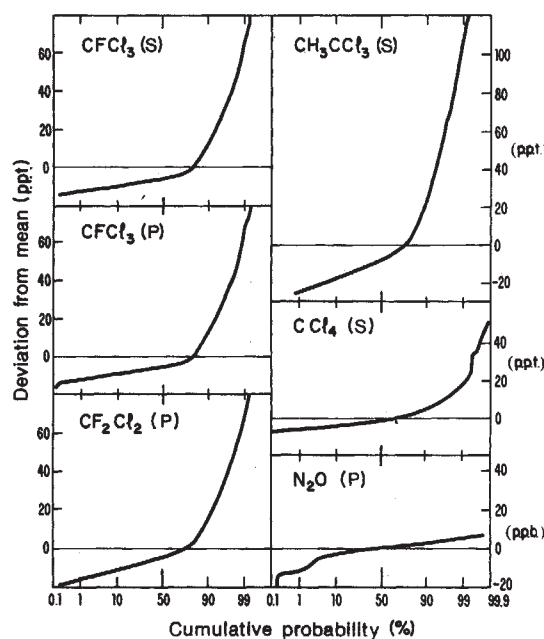


Fig. 2 Probability-distribution functions for the ALE data from Adrigole, Ireland. The primary data shown in Fig. 1 have been sorted into intervals of 8 h, and a least-squares, linear trend has been removed.

Table 1 ALE data from Adrigole

Gas	Column	Factor	ALE analysis		Present analysis	
			Mean	Trend (% yr ⁻¹)	Mean truncated/full	Trend (% yr ⁻¹)
CFC-11	P	0.96	178.8	+5.1 ± 0.2	184.6/184.9	+4.8 ± 0.3
CFC-12	P	0.95	301.3	+6.3 ± 0.2	307.1/307.4	+6.0 ± 0.2
N ₂ O	P				306.0/306.0	-0.2 ± 0.5
CFC-11	S	0.96	177.9	+5.0 ± 0.2	184.2/184.5	+4.4 ± 0.3
CH ₃ CCl ₃	S	0.80	152.3	+7.7 ± 0.6	161.6/161.9	+7.9 ± 0.7
CCl ₄	S	0.80	124.6	+2.7 ± 0.6	126.2/126.4	+2.8 ± 0.4

Mean values refer to January 1980 and have units of parts per 10¹² by volume, except for N₂O with units of parts per 10⁹ by volume. Suggested calibration factors have not been applied (see ref. 9). ALE analysis does not include the 'pollution events', ~30% of observations. Reported trends are for the truncated data set; those for the full set differ by less than the quoted error (95% confidence range for the linear fit).

Table 2 Coherences with respect to CFC-11S

±lag (h)	CFC-11P	CFC-12	N ₂ O	CH ₃ CCl ₃	CCl ₄
0	0.92	0.89	0.19	0.86	0.56
8-24	1.00	0.97	0.28	0.94	0.71
32-104	1.01	0.98	0.35	0.99	0.90
112-176	1.00	0.97	0.32	0.95	1.08

Values are based on the initial (linear-trend) analysis described in text.

Fig. 1). Autocorrelations for the six ALE gases are shown in Fig. 3; and selected cross-covariances, in Figs 4 and 5.

Trace gas correlations

All gases with the exception of N₂O (that is, the halocarbons) exhibit similar patterns: the full width of a pollution event is 2 days at half height; enhanced concentrations persist for long periods (autocorrelations of 20% peak variance extend out to ±5 days); a significant repetition of these events occurs on a scale of 5-8 days with additional structure out to 30 days. Air containing elevated concentrations of halocarbons (such as, variances of 200 (parts per 10¹² by volume, p.p.t.)² is clearly identifiable against the mean background (such as, variances of 10 (p.p.t.)² derived from the probability distribution in Fig. 2). Both N₂O and CCl₄ data are significantly, positively correlated at ±30 days, a consequence of seasonal/annual variability not removed by a linear trend, or possibly a result of their relatively shorter data records (see Fig. 1). The different halocarbons are indeed highly correlated with each other as may be seen in Fig. 4, where the cross-covariances show the same patterns evidenced in the autocorrelations.

For N₂O the uncorrelated noise in the autocorrelation at a time lag of zero, presumably the instrumental variance, is much greater than the correlated variance for lags greater than zero. Its magnitude, ~5 (parts per 10⁹ by volume, p.p.b.)², is consistent with that derived from the slope of the probability distribution, ~7 (p.p.b.)². Furthermore, the N₂O observations show a broad autocorrelation of the order of 1.5 p.p.b., extending over a range of ±15 days. Examination of the N₂O observations in Fig. 1 points to systematic shifts in concentration every calendar month that might lead to such a correlation extending over 30 days. Because of these problems with N₂O, I have re-examined the original data with a different analysis, replacing steps (2)-(6) with the following:

- (2') Reference values for the 8-h data are selected on a month-by-month basis, setting the baseline at the 50th percentile (that is, median) of observations for the month.
- (3') The data are sorted as before.
- (4') The ±0.5% extreme values are truncated as before.
- (5') In a further effort to reduce noise, the apparently aberrant

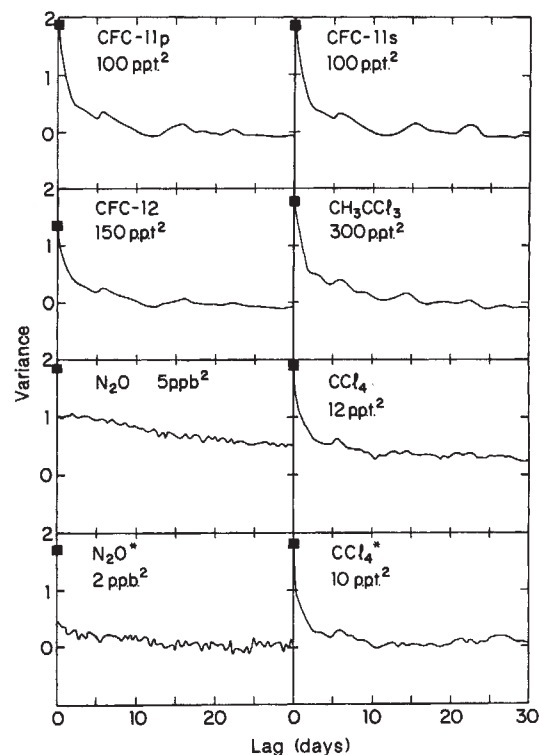


Fig. 3 Autocorrelations of the six ALE gases from Adrigole. The filled square denotes the zero-lag autocorrelation (white noise) of the time series. Scale factors for the variances are given in each panel. The upper six panels refer to time series from which the linear trend has been removed; the lower two panels (*) show results from the monthly median analysis.

N₂O observations from March through June of 1980 have been dropped (see Fig. 1).

(6') Correlations and covariances are computed as before; selected correlations from this monthly median analysis are included in Figs 3 and 5, denoted by asterisks.

This second analysis does not significantly alter the halocarbon self-correlations, but it greatly reduces the noise in the N₂O time series. As expected, when the data are renormalized by monthly medians, autocorrelations at ±30 days are nearly zero. The covariance of N₂O with the other halocarbons is now more readily visible, and the uncorrelated noise in the data has been reduced to 2.5 (p.p.b.)². The N₂O signal which correlates with pollution events is small, <0.5 p.p.b., and care is needed to ensure that it is not caused by interference in the gas chromatograph with a pollution-associated trace gas such as CO₂. The coherence of N₂O with respect to CFC-11S now exceeds 0.4 at 5-8-day lags.

No pair of gases is perfectly coherent at zero lag, not even CFC-11S and CFC-11P. This may be understood in that a time lag of zero does not mean that both S and P columns sampled the same air mass, but rather that both measurements were made within 8 h of each other. Beginning at a lag of ±8 h and extending out to ±8 days the major halocarbons (CFC-11S, CFC-11P, CFC-12 and CH₃CCl₃) become highly coherent (0.98 ± 0.03), indicating that pollution events of duration >8 h have evolved from quantitatively comparable sources of these gases. Carbon tetrachloride, on the other hand, is not fully coherent with the other halocarbons (for example, 0.71 at 8-24 h) until we reach lags of 5-8 days. The asymmetry of the CCl₄ cross-covariances would imply that elevated levels of CCl₄ are likely to precede a general halocarbon pollution event by about 10 days. A preliminary interpretation of this pattern would be that the major CCl₄ sources observed at Adrigole are indeed European but of spatially distinct regions (for example, eastern Europe) as compared with those of major CFC sources (for example, western Europe).

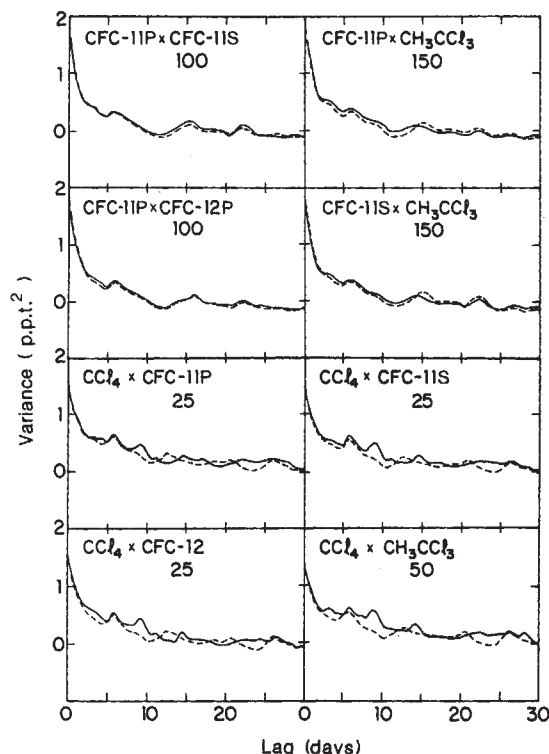


Fig. 4 Cross-covariances of the halocarbons at Adrigole. The covariances have been defined in the sense given by the formula in the text. Each panel denotes the cross of $f_1 \times f_2$ with the solid line referring to positive lags ($+\Delta$) and the dashed line to negative lags ($-\Delta$). Scale factors for the variances are included in each panel.

Sources

European source strengths for CFC-12, CH_3CCl_3 , CCl_4 and N_2O with respect to that for CFC-11 are derived by two methods and reported in Table 3. We may determine relative emissions from the probability distributions if the highest concentrations observed for each gas are assumed to be perfectly correlated, such that the difference in concentrations between, for example, the 90 and 98% levels is proportional to the source. Molar ratios of these inferred sources are calculated from the probability distributions in Fig. 2 with the background (variance) removed. Relative sources of trace gas 2 with respect to reference gas 1 are derived from the time-series by averaging the ratio of the relative covariance ($c_{12}(\Delta)$) to the autocorrelations of the refer-

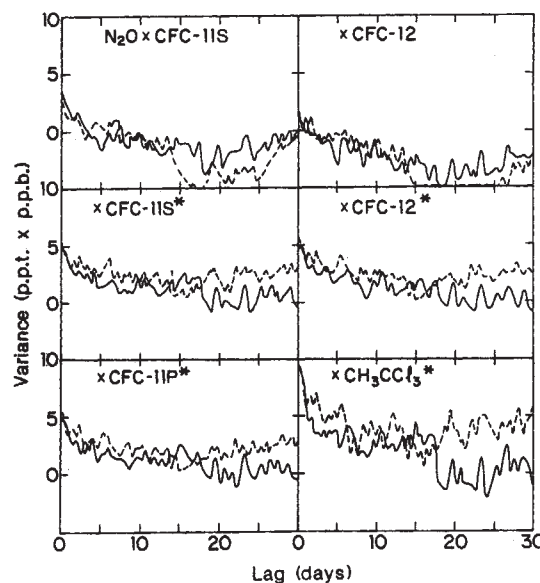


Fig. 5 Cross-covariances of N_2O with halocarbons at Adrigole. The covariances are defined as in Fig. 4. The upper panels refer to the analysis with linear trends removed; the lower four panels (*) are based on the monthly median analysis with some of the N_2O data excluded (see text).

ence gas ($c_{11}(\Delta)$). Averages are based on time lags >8 h and <8 days; contemporaneous measurements (zero lag) are not used because their coherence is low and the instrumental noise may be greater than the atmospheric signal (for example, N_2O at Adrigole, CFCs at Cape Meares).

Nitrous oxide sources have been estimated with respect to all the halocarbons and span more than a factor of 2 as indicated. Because the halocarbons are so highly coherent, estimates based on the probability distributions alone agree with the more accurate time-series analyses. Table 3 also includes results of similar analyses for the ALE data from Cape Meares, Oregon and Cape Grim, Tasmania. These latter two stations observe pollution events of smaller magnitude less frequently than Adrigole, requiring a time-series analysis to characterize accurately the atmospheric events above instrumental noise. Unfortunately, no estimates of N_2O sources may be inferred from the Oregon or Tasmania data.

These derived emissions are independent of the local meteorology and synoptic processes as long as the variability at Adrigole results from addition of these inert tracers to a relatively well mixed, uniform background atmosphere. Meteorological studies of halocarbon pollutions at Adrigole^{9,10} clearly point to air that has travelled over the European continent and the British Isles. An absolute calibration of the halocarbon emissions would require not only knowledge of meteorological conditions, but also a proven model for boundary layer mixing and atmospheric transport. Nevertheless, case studies of the CCl_4 episodes may enable positive identification of this source with eastern Europe.

Conclusions

Release patterns for the CFCs are similar for Europe and the western United States, although Australian releases of CFC-12 with respect to CFC-11 are noticeably smaller. Most importantly, global budgets for CFCs^{9,12}, based on observed atmospheric trends and on industrial inventories, indicate a substantially larger ratio of CFC-12 emission to CFC-11 emission (~ 1.8) than is observed for all three regions studied here (~ 1.1). The Soviet Union and developing countries have been proposed as playing a relatively more important role in world use of CFC-12. The Chemical Manufacturers Association (CMA) estimates¹² of the CFC-12: CFC-11 ratio for reporting (western) companies (~ 1.5)

Table 3 Relative molar source strengths

	CFC-11	CFC-12	CH_3CCl_3	CCl_4	N_2O
Probability distribution*					
ALE (Ireland)	1	1.0 ± 0.1	1.9 ± 0.2	0.23 ± 0.07	-0 ± 100
Time series†					
ALE (Ireland)	1	1.1	1.8	0.28	30–70
ALE (Oregon)	1	1.2 ± 0.2	2.5 ± 0.3	0.15 ± 0.05	ND
ALE (Tasmania)	1	0.9 ± 0.2	0.9 ± 0.2	0.15 ± 0.04	ND
Industrial emissions‡					
Global (flux in 10^9 yr^{-1})	1 (265)	1.76 (410)	1.94 (500)	0.34 (100)	52 (2,800 of N)

* Dimensionless ratios are based on the probability distribution between the 90th and 98th percentile ranking; error estimates are subjective, including uncertainties in determination of the high-end probability distributions.

† Error estimates are $\pm 10\%$ for all but N_2O (see text). Absolute calibration factors for the ALE data have not been applied, and would reduce the relative sources for CCl_4 and CH_3CCl_3 by $\sim 15\%$. N_2O sources for Oregon and Tasmania are not detectable (ND).

‡ Estimates for industrial emissions are generally available only for the entire Northern Hemisphere (see refs 9, 12). N_2O flux estimate corresponds to $0.2\% \text{ yr}^{-1}$.

is lower than the global average but still greater than observed for the United States and Europe.

Emissions of CH_3CCl_3 (with respect to CFC-11) are clearly different from the three continents: the western United States is a factor of 1.3 above the global mean, and Australia a factor of 2 below. A balance of sources from the United States and western Europe could maintain the CH_3CCl_3 global budget.

Carbon tetrachloride shows a large European source (with respect to CFC-11) which is still less than estimates for the mean global source. As the United States and Australia seem to contribute lesser relative amounts of CCl_4 , I conclude that additional sources may be required, perhaps also associated with the missing CFC-12. The primary pollution events at Adrigole show a CCl_4 :CFC-11 ratio typical of the other continents (~ 0.15). The larger sources of CCl_4 are associated with high coherence at time lags or more than 4 days, enough time for air to have crossed the European continent.

Nitrous oxide variability at Adrigole shows a component clearly correlated with halocarbon pollution events. This

relationship is consistent with an industrial source of N_2O from combustion of fuel nitrogen^{13,14} which is spatially associated with halocarbon releases. If scaled to global CFC-11 emissions, the magnitude of this source is $\sim 3 \times 10^{12} \text{ g(N) yr}^{-1}$, is equivalent to the observed annual increase in N_2O , and accounts for $\sim 25\%$ of the current global emissions⁷.

Concurrent atmospheric measurements, made at a frequency greater than one per day, provide valuable information on the location and magnitude of sources for many trace gases of both natural and anthropogenic origin. These data permit integration of emissions over large areas and present critical tests for the synoptic-scale transport of trace gases in three-dimensional models. Locations of observing stations near, but not in the middle of suspected sources would allow determination of baseline trends¹⁷ as well as emissions.

This research was supported by NSF grant ATM-81-17009 and NASA grant NAGW-359. I gratefully acknowledge availability of the atmospheric measurements from the ALE team and helpful discussions with S. Wofsy and R. Prinn.

Received 28 March; accepted 10 July 1985.

1. Keeling, C. D. *et al.* *Tellus* **28**, 538–551 (1976).
2. Dianov-Klovov, V. I. & Yurganov, L. N. *Tellus* **33**, 262–273 (1981).
3. Rasmussen, R. A. & Khalil, M. A. K. *Atmos. Environ.* **15**, 883–886 (1981).
4. Ehhalt, D. H., Zander, R. J. & Lamontagne, R. A. *J. geophys. Res.* **88**, 8442–8446 (1983).
5. Craig, H. & Chou, C. C. *Geophys. Res. Lett.* **9**, 477–481 (1982).
6. Khalil, M. A. K. & Rasmussen, R. A. *Science* **224**, 54–56 (1984).
7. Weiss, R. F. *J. geophys. Res.* **86**, 7185–7196 (1981).
8. Fraser, P. J., Hyson, P., Enting, I. G. & Perman, G. I. *Nature* **302**, 692–695 (1983).
9. Prinn, R. G. *et al.* *J. geophys. Res.* **88**, 8353–8441 (1983).

10. *Causes and Effects of Changes in Stratospheric Ozone: Update 1983* (National Research Council, National Academy, Washington DC, 1984).
11. Prather, M. J., McElroy, M. B. & Wofsy, S. C. *Nature* **312**, 227–231 (1984).
12. *Production, Sales and Calculated Release of CFC-11 and CFC-12 Through 1982* (Chemical Manufacturers Association, Washington DC, 1983).
13. Weiss, R. F. & Craig, H. *Geophys. Res. Lett.* **3**, 751–753 (1976).
14. Pierotti, D. & Rasmussen, R. A. *Geophys. Res. Lett.* **3**, 265–267 (1976).
15. Blake, D. R., Woo, V. H., Tyler, S. C. & Rowland, F. S. *Geophys. Res. Lett.* **11**, 1211–1214 (1984).
16. Pack, D. H., Lovelock, J. E., Cotton, G. & Curthoys, C. *Atmos. Environ.* **11**, 329–344 (1977).
17. Graedel, T. E. & McRae, J. E. *Geophys. Res. Lett.* **11**, 977–979 (1980).

Segmentation of mid-ocean ridges

Hans Schouten*, Kim D. Klitgord† & John A. Whitehead*

* Woods Hole Oceanographic Institution, Woods Hole, Massachusetts 02543, USA

† US Geological Survey, Woods Hole, Massachusetts 02543, USA

Studies of mid-ocean ridges in the Pacific and Atlantic oceans show that the volcanism that forms the oceanic crust along the spreading-plate boundaries is concentrated at regular intervals related to spreading rate. This observation and a new calculation for a Rayleigh–Taylor type of gravitational instability of a partially molten mantle region growing under spreading centres yield reasonable estimates of upper mantle viscosities.

SEGMENTATION of mid-ocean ridges seems to be a global phenomenon. Mid-ocean ridges are the spreading boundaries or spreading centres between lithospheric plates where new oceanic crust and lithosphere is created at rates of 10–200 km Myr^{-1} . Although continuous overall, mid-ocean ridges are cut by many transform zones and associated fracture zones that offset the spreading centre by distances up to several hundred kilometres (Fig. 1). These fracture zones divide the ridges into a series of relatively straight segments with lengths of several tens to several hundreds of kilometres. As more high-resolution geophysical surveys of mid-ocean ridges are becoming available in the world's ocean basins, an even more orderly pattern of the sea floor is emerging. This pattern indicates that the boundaries between the spreading plates have a fairly rigorous cellular structure on a 30–80 km length scale.

The spreading centre cells seem to be the primary feature of mid-ocean ridges; they suggest a three-dimensional, rather than two-dimensional, process of magmatism and crustal accretion spaced evenly along the present spreading boundary. Magnetic studies of older oceanic crust suggest relatively stable and even segmentation of plate boundaries in the past on the same length scale^{1,2}. Seismic studies indicate that oceanic crust with normal seismic structure is formed in the spreading centre cells and that between the cells, thinner crust and other structural anomalies

characteristic of fracture zones are formed, regardless of the spatial offset between the spreading centres^{3–7}.

The transition zones or transform zones between centres on the active mid-ocean ridges manifest themselves in variable and complex ways. Where large offsets exist between the centres, well-defined, predictable transform fault zones provide the continuity and stability of the active plate boundary between the spreading centres on either side of the transform. Where no offsets are apparent between the spreading centre cells, the transition zones between centres are complex, variable and inherently unstable on an approximate length scale of 10 km and timescale of 10^5 – 10^6 yr (refs 8, 9), but they are apparently stable when observed on longer length- and timescales^{1–3,7}.

The apparent stability in space and time on the longer length- and timescales of a regular segmentation of mid-ocean ridges has been attributed by Whitehead, Dick and Schouten¹⁰ to a Rayleigh–Taylor type of gravitational instability of a partially molten mantle region beneath spreading centres. In this model, the wavelength of the instability controls the aggregation and diapiric ascent of mantle melt (magma) at regular intervals along the boundary between two spreading plates.

Here we use the results of high-resolution geophysical surveys of mid-ocean ridges and determine the length scales of segmentation as a function of spreading rate. We use a modification of

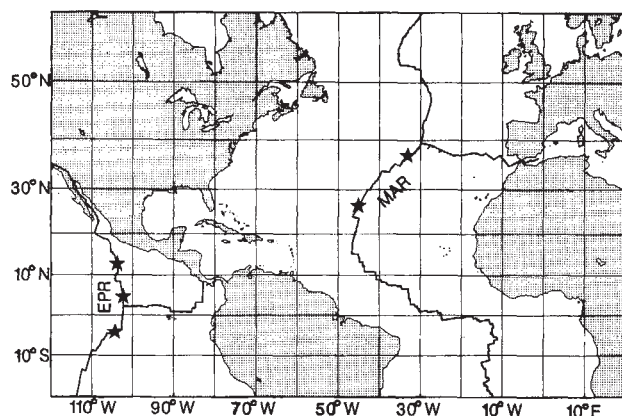


Fig. 1 Atlantic and Pacific spreading centres and transform zones which form the boundaries between the spreading lithospheric plates. Stars denote the location of five high-resolution surveys of the Mid-Atlantic Ridge (MAR) and East Pacific Rise (EPR) discussed in this article.

a theory by Howard¹¹ for Rayleigh–Benard convection to predict length- and timescales of magmatic activity as a function of spreading rate. We show that observed length- and timescales of magmatic activity on mid-ocean ridges predict reasonable upper mantle viscosities, suggesting that the chosen length- and timescales and the proposed theory are appropriate. This provides additional supporting evidence for the mechanism for magmatic accretion under spreading centres proposed by Whitehead *et al.*¹⁰ and for the predictability and inherent stability of length- and timescales of magmatism and crustal accretion on past and present-day mid-ocean ridges.

Spreading centre segmentation

A detailed SEABEAM investigation of the East Pacific Rise (EPR in Fig. 1), which provided continuous along-strike coverage of the rise-axis bathymetry between 8° N and 18° N, revealed that the continuity of a 550-km-long spreading-centre segment between the Clipperton Fracture Zone (10° N) and the Orozco Fracture Zone (15° N) was disrupted in at least six locations by a volcano-tectonic geometry characterized by two overlapping (en-echelon) neovolcanic zones enclosing an elongate depression^{8,9} (EPR 10°–14° N in Fig. 2). The existence of similar disruptions of the southern East Pacific Rise axis between 20° S and 7° N has been reported by Lonsdale^{12–14} (EPR 3°–6° N in Fig. 2). Schouten and Klitgord² have argued that these geometries are the morphological expression on fast spreading ridges of ‘zero-offset’ transform zones between spreading-centre cells which have been stationary features on the EPR spreading axis for at least 5 Myr.

The segmentation of the mid-Atlantic Ridge (MAR in Fig. 1) has been suggested by interpretations of high-resolution surveys in the FAMOUS area¹⁵ (MAR 36°–37° N in Fig. 2) and in the TAG area^{16,17} (MAR 25°–27° N in Fig. 2) which indicate a spreading-centre spacing of the order of 40 and 45 km, respectively. Any spatial offset between two rift valleys translates into a significant age offset on this slowly spreading ($\sim 25 \text{ km Myr}^{-1}$) plate boundary. Age offset rather than spatial offset seems to characterize the anomalous great depths found at ridge/transform intersections¹⁸, causing significant disruptions in the rift-valley bathymetry. If we interpret each rift-valley segment bounded by anomalous depths as a spreading-centre cell, then bathymetry between the Kane Fracture Zone at 24° N and the Azores triple junction of the North American, African and European plates at 38° N would give an estimated distance between cell centres of $50 \pm 10 \text{ km}$, which is in agreement with the high-resolution observations in FAMOUS and TAG. On the faster spreading EPR we interpret the disruptions of the neovolcanic zones shown in Fig. 2 as the boundaries between

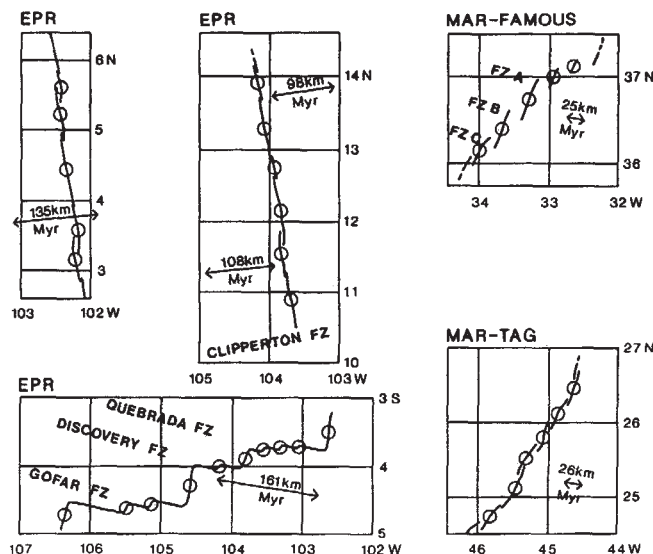


Fig. 2 Segmentation of seafloor spreading centres from high-resolution geophysical data on EPR and MAR. Lines denote neovolcanic zones inferred from ridge-axis topographical highs (on fast-spreading EPR) and deeps (on slow-spreading MAR). EPR 3°–6° N after Lonsdale¹⁴ (multi-narrow-beam bathymetry data); EPR 10°–14° N after Macdonald and Fox⁸ and Macdonald *et al.*⁹ (multi-narrow-beam bathymetry data); MAR 25°–27° N modified from Rona *et al.*¹⁶ and Rona and Gray¹⁷ (single narrow-beam bathymetry data, average track spacing 7 nm); MAR 35°–37° N modified from Ramberg and van Andel¹⁵ (multi-narrow-beam and single-beam bathymetry data). The heavy line in EPR 5°–3° S denotes the Pacific–Nazca plate boundary after Searle³⁰ (GLORIA long-range side-scan sonar data). Open circles, placed at midpoints of the spreading-centre segments, denote hypothetical centres of magmatic activity. Arrows indicate total seafloor spreading (after ref. 38). Average spacing of these centres as a function of effective spreading rate is shown in Fig. 3. We postulate that the fairly even spacing of these centres can be attributed to a Rayleigh–Taylor type of gravitational instability of a lower density and viscosity region of high-melt content collecting at the base of melt-depleted mantle at some depth beneath the spreading plate boundary. The reduced distance between centres in EPR 5°–3° S can be attributed to reduced rate of melt production per unit length along this oblique section of the spreading plate boundary; the average effective spreading rate on this oblique boundary is of the order of 75 km Myr^{-1} or less than half the spreading rate at this location.

spreading-centre cells, indicating a fairly even spreading-centre spacing of the order of 65 km.

Three-dimensional magmatic accretion

Submersible observations and bottom camera surveys at spreading centre/transform zone intersections indicate an apparent reduction of magmatic activity towards the transform zone regardless of transform offset (see refs 18, 19). Near-bottom observations of the Galapagos Rift at 86° W have shown clear along-strike variations in volcanism, structure and hydrothermal activity along a 30-km section of a 250-km long spreading-centre segment, suggesting a three-dimensional rather than two-dimensional process of magmatism and crustal accretion²⁰. Similar observations of the along-strike variability of axial processes on the EPR near 20° S, 13° N and 21° N and on the Juan de Fuca Ridge support a three-dimensional model of magmatic accretion in seafloor spreading centres^{21,22}.

Segmented magmatic accretion

The cellular structure of the boundaries between spreading plates is one of the arguments used by Whitehead *et al.*¹⁰ and Crane²³ to support a new model of magmatic accretion under spreading centres. In this model, the mechanism that controls the segmentation of mid-ocean ridges is attributed to a Rayleigh–

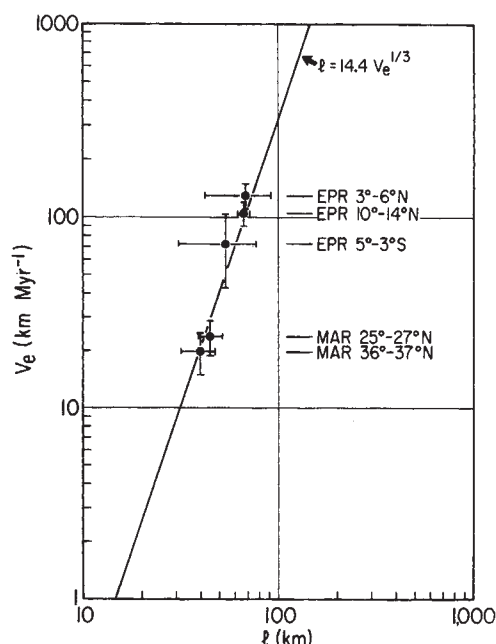


Fig. 3 Spacing (l) of hypothetical centres of magmatic activity (from Fig. 2) against effective spreading rate v_e (error bars are ± 1 s.d.) and a best-fit model for the proposed relationship where $l \propto v_e^{1/3}$, or $l = 14.4 v_e^{1/3}$. This model predicts an average spreading-centre spacing of 85 km for a 200 km Myr^{-1} spreading rate at 20° S on the EPR.

Taylor type of gravitational instability of a lower viscosity and lower-density partially molten region collecting at the base of melt-depleted mantle at some depth within the asthenosphere rising beneath the spreading-centre axis. The wavelength or spacing of the fastest growing instability of the low-viscosity material is a function of the geometry of that region and of the viscosities of the partially molten region and the surrounding asthenosphere^{10,24,25}.

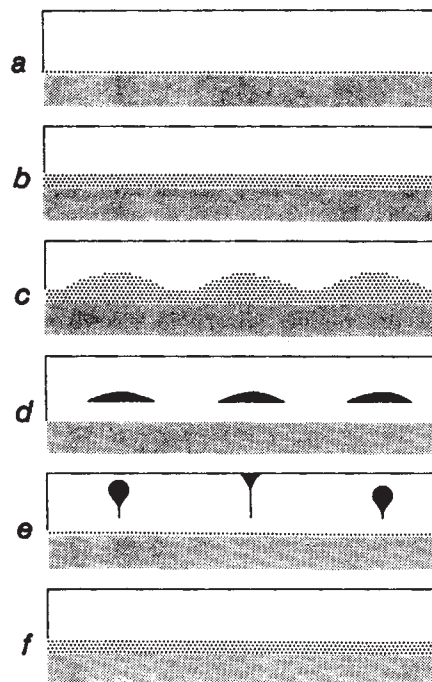
Considering that mid-ocean ridge basalts (the mantle melt that has aggregated from a partially molten mantle and risen to the surface to form the oceanic crust) have a rather characteristic and uniform composition throughout the world²⁶, it is probable

that the viscosities of the partially molten region and the surrounding asthenosphere beneath mid-ocean ridges are uniform also, and independent of spreading rate. However, the production of melt, caused by the decompression of the asthenosphere as it rises to fill the gap between two spreading plates, must at least equal production of oceanic crust. If we make a first-order approximation that the thickness of oceanic crust is uniform and independent of spreading rate, then production of melt beneath mid-ocean ridges would be proportional to spreading rate.

The model of Whitehead *et al.*¹⁰ postulated a linearly continuous partially molten region beneath the spreading plate boundary. The width of this region cannot exceed the width of the asthenosphere rising between the spreading lithospheric plates. This width is a function of depth or age as the lithospheric plates thicken with age away from the axis^{27,28}. At 25 km depth beneath the spreading axis, where the instability is postulated to occur¹⁰, the width of the asthenosphere between the lithospheric plates is the total spreading rate multiplied by 4 Myr (ref. 27) to 7 Myr (ref. 28). On the fast-spreading EPR (100–200 km Myr^{-1}), between Rivera Fracture Zone (18° N) and Eltanin Fracture Zone (57° S), most transform offsets are ≤ 1 Myr times the spreading rate (comparable with offsets of ≤ 25 km in the slow-spreading central North Atlantic). These offsets would be insufficient to disrupt the continuity of a partially molten region at 25 km depth beneath the overall spreading boundary. A linearly continuous partially molten region beneath the EPR axis, with its gravitational instability driving the segmentation of this axis, would explain the phenomenon of multiple closely spaced transform zones. The double Orozco Fracture Zone at 15° N (ref. 29) provides the continuity of regular spacing (of the order of 65 km) between centres of magmatic activity along an oblique part of the overall plate boundary. The multiple Quebrada, Discovery and Gofar Fracture Zones near 5° S on the EPR surveyed with GLORIA³⁰ (EPR 5°–3° S in Fig. 2) suggests a fairly regular spacing between centres of magmatic activity (of the order of 55 km) along a very oblique section of the overall spreading boundary (the trend is $\sim 30^\circ$ to the spreading direction).

To include the segmentation of oblique plate boundaries in our estimates, we postulate that the average production of melt per unit length along the average boundary is proportional to the effective spreading rate (where effective spreading rate is the component of spreading perpendicular to the regional trend of the spreading plate boundary). A plot of average distance

Fig. 4 Along-axis diagram of the proposed mechanism of magmatic upwelling beneath mid-ocean spreading centres. Fine stipples, parental upper mantle; coarse stipples, partially molten mantle; no stipples, melt-depleted mantle; solid black, aggregated mantle melt (magma). A region of partially molten mantle begins to collect at the base of the depleted mantle in the asthenosphere beneath the spreading axis (a). As the volume of this region grows (b), the partially molten mantle, which has a lower density and viscosity than the surrounding asthenosphere, develops a gravitational instability with regularly spaced disturbances (c). Once the disturbances are established, the very low viscosity melt (magma), which can migrate by porous flow and moves more rapidly than the partially molten region as a whole, concentrates at the tops of the rising disturbances (d), after which it rises diapirically (e) towards the magma chambers in the oceanic crust above. The partially molten region is depleted of the melt (d) and the process will repeat itself as a new region of partially molten mantle begins to grow again at the base of the freshly depleted mantle (e, f).



between centres of magmatic activity (that is, the centres of spreading-centre segments indicated by circles in Fig. 2) versus effective spreading rate is shown in Fig. 3. The segmentation of mid-ocean ridges would be proportional approximately to one-third power of the effective spreading rate (Fig. 3).

Rayleigh–Taylor instability

We use a modification of a remarkably simple theory by Howard¹¹ for Rayleigh–Benard convection (a horizontal layer of fluid heated from below) to predict spacing between centres of magmatic activity as a function of spreading rate. In Howard's theory, the conductive growth of a thermal boundary layer from an initially isothermal state was shown to proceed as an error function. The boundary layer could be next to either the cold top plate or the hot bottom plate. He postulated that the boundary layer would grow until its local Rayleigh number reached a number of 1,000–3,000. Next, a disturbance would grow more rapidly than the conductive boundary layer, and a blob of fluid would detach from the boundary. The boundary layer would be swept free of the boundary and the process would repeat itself as fluid near the boundary began to form a new thermal boundary layer. The calculation led to the correct prediction that heat transport would be proportional to Rayleigh number (dimensionless temperature difference) to the one-third power, a result known to be true both from experiment and by more rigorous and much more complicated theories^{11,31,32}.

Here we do not have a conductive boundary layer but instead consider a growing horizontal cylindrical body of fluid with relatively low viscosity and density compared with the surrounding liquid. This represents a linear region of relatively high melt content growing in the Earth's mantle below the spreading boundary. To be consistent with uniform formation of oceanic crust, we postulate that the volume of the cylindrical body increases uniformly with time. The diameter of this body will increase proportionally to the square root of time, that is

$$d = \left(\frac{4Ut}{\pi} \right)^{1/2} \quad (1)$$

where U is the two-dimensional volumetric production rate of the lower viscosity and density region and t is time. We imagine this is the first part of a cycle in which upwelling mantle develops a linear region of lower viscosity and of density $\rho - \Delta\rho$.

The appropriate gravitational instability for a horizontal cylindrical volume is a modified Rayleigh–Taylor instability^{10,25} for which it has been shown that small disturbances will grow exponentially as $\exp(t/\tau)$, where (using equation (1)) the time-scale τ is

$$\tau = \frac{1}{0.153} \frac{\mu_2}{g\Delta\rho} \left(\frac{\pi}{4Ut} \right)^{1/2} \left(\frac{\mu_2}{\mu_1} \right)^{-1/3} \quad (2)$$

Here μ_2 is the viscosity of the surrounding mantle fluid, μ_1 the viscosity of the cylindrical volume of high-melt content and g is the force of gravity. Although this equation is strictly valid only for large values of μ_2/μ_1 , it is quite accurate for values as low as 1.5 or 2. Equation (2) shows that τ is initially infinite, but that it decreases inversely as the square root of time (that is, diameter). As a modification to Howard's theory, let us assume that as the volume grows, no disturbances will be apparent until a time (t_c) of the order of τ (Marsh²⁵ advocates a factor of 5τ before protrusions are seen experimentally). Setting $t = t_c = \tau$ in equation (2) and solving for t_c , we get

$$t_c = \left[\frac{1}{0.173} \frac{\mu_2}{g\Delta\rho} \left(\frac{1}{U} \right)^{1/2} \left(\frac{\mu_2}{\mu_1} \right)^{-1/3} \right]^{2/3} \quad (3)$$

The wavenumber of fastest growing disturbances for a cylindrical volume is advocated by Marsh²⁵ to be

$$k_c = \frac{2.15}{d} \left(\frac{\mu_1}{\mu_2} \right)^{1/3} \quad (4)$$

Using equations (1), (3) and (4) and $k_c = 2\pi/\lambda_c$, where λ_c is the critical wavelength, we get the following expression for the

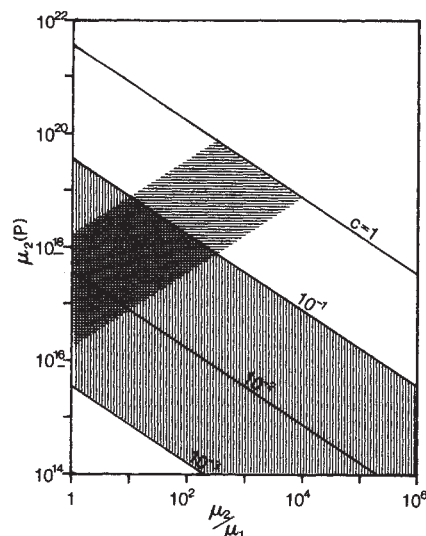


Fig. 5 Constraints on upper mantle viscosity (μ_2) and the ratio, μ_2/μ_1 , of mantle viscosity to viscosity of the partially molten mantle region (μ_1) under seafloor spreading centres. c , Concentration of melt in the partially molten region. Lines of constant c are drawn from equation (9) which relates the wavelength of a Rayleigh–Taylor instability to observed length scales of magmatic activity on mid-ocean ridges. Values of $c > 1$ are impossible, and $(\mu_2/\mu_1) < 1$ is considered to be physically unlikely. Vertical hachures indicate a permissible range of melt concentrations $10^{-1} > c > 10^{-3}$; greater values would result in semi-suspended slurry with too-large values of μ_2/μ_1 ; smaller values of c would not result in a sizeable magma genesis. A second constraint (horizontal hachures) arises from the observed time- and length scale of magmatic activity in the central North Atlantic FAMOUS area and from equation (6). The centre of intersecting hachures lies at $\sim \mu_2 = 10^{18}$ P, $(\mu_2/\mu_1) = 10$, $c = 3\%$. The fact that these numbers are reasonable suggests that length- and timescales of magmatic activity on mid-ocean ridges can provide important boundary conditions for models of mantle upwelling under spreading centres.

wavelength of fastest-growing disturbances at critical time t_c , that is

$$\lambda_c = \frac{1}{0.169} U^{1/3} \left(\frac{\mu_2}{g\Delta\rho} \right)^{1/3} \left(\frac{\mu_2}{\mu_1} \right)^{2/9} \quad (5)$$

We suggest that once disturbances with this spacing are established in the partially molten mantle region, the very low viscosity melt, which can migrate by porous flow³³ would move more rapidly than the partially molten region as a whole and concentrate at the tops of the rising disturbances, after which it would be drawn off through conduits¹⁰ (Fig. 4). The partially molten region would be depleted of the melt and the process would repeat itself as a new region of high-melt content begins to grow at the base of the freshly depleted mantle.

Of special interest will be the product of equations (3) and (5), which reduces to the simple expression

$$\lambda_c t_c = 19.1 \frac{\mu_2}{g\Delta\rho} \quad (6)$$

Discussion

The most striking property of equation (5) is the dependence of spacing on production rate to the one-third power, which matches the empirical dependence found in Fig. 3 of spreading-centre spacing l on effective spreading rate v_e to approximately the one-third power. We postulate a model of the form of

$$l = Q v_e^{1/3} \quad (7)$$

to fit the data in Fig. 3, and choose $Q = 14.4 \text{ km}^{2/3} \text{ Myr}^{1/3}$ as a best fit. We relate the volumetric production rate U of a partially

molten mantle region beneath spreading centres to the three quantities: effective spreading rate, thickness of oceanic crust T and the melt concentration c in the partially molten region. We assume that production of melt equals production of oceanic crust, thus the production rate of the partially molten region is given by

$$U = \frac{v_c T}{c} \quad (8)$$

Equations (5), (7) and (8) can be combined to relate the empirical parameter Q with μ_1 , μ_2 , $\Delta\rho$, c and T . Setting $l = \lambda_c$, we obtain

$$\mu_2 = (0.169 Q)^3 \frac{c g \Delta\rho}{T} \left(\frac{\mu_2}{\mu_1} \right)^{-2/3} \quad (9)$$

This relationship is illustrated in Fig. 5 with the following numbers for the observed quantities: $Q = 0.98 \times 10^9 \text{ cm}^{2/3} \text{ s}^{1/3}$, average thickness of oceanic crust (fracture zones included) $T = 5 \times 10^5 \text{ cm}$ and $g = 10^3 \text{ cm s}^{-2}$. It is further assumed that $\Delta\rho$ of the partially molten region is proportional to melt concentration, that is, $\Delta\rho = 0.4 c \text{ gm cm}^{-3}$. Thus, our remaining unknowns are μ_2 , μ_2/μ_1 and c ; lines of constant c are presented in a μ_2 , μ_2/μ_1 plane. Vertical hachures in Fig. 5 indicate a permissible range of melt concentrations $10^{-1} > c > 10^{-3}$. If the postulated dynamics are correct and if the value of Q adopted here predicts the correct spreading-centre spacing on the world's mid-ocean ridges, then Fig. 5 will provide important constraints on the viscosities of the upper mantle and the partially molten region below spreading centres.

An additional physical prediction of the model is the critical time t_c , which can be thought of as a timescale of magmatic activity. From equations (3) and (8) it is easy to see that also the frequency ($1/t_c$) of the process should be proportional to effective spreading rate to the one-third power, predicting a frequency of magmatic activity in the fastest spreading centres (EPR 20° S at 200 km Myr^{-1}) that is only a factor of two greater than the frequency of magmatic activity in slow-spreading centres (MAR at 25 km Myr^{-1}). Recent near-bottom studies of spreading-centre activity along the EPR axis indicate a frequency of magmatic activity which is only a factor of two or three higher than that of the FAMOUS area and Reykjanes Ridge on the slow-spreading MAR (R. D. Ballard, personal communication) which is consistent with the prediction from our model.

From an interpretation of FAMOUS rift valley geomorphology, Ballard and van Andel³⁴ estimate a period of magmatic activity of the order of 10^4 yr . The wavelength of magmatic activity or spreading-centre spacing in this area is 40 km (MAR 36°–37° N in Fig. 2). We use these estimates of period and wavelength from the FAMOUS area with equation (6) to predict

upper mantle viscosity μ_2 . Taking $t_c = 10^{4 \pm 0.5} \text{ yr}$, $\lambda_c = 40 \text{ km}$, $g = 10^3 \text{ cm s}^{-2}$ and assuming $\Delta\rho = 0.4 c \text{ g cm}^{-3}$ we find from equation (6) that $\mu_2 = 2.64 \times 10^{19 \pm 0.5} \text{ c}$. This estimate is represented with the horizontal hachures in Fig. 5. The area of intersecting vertical and horizontal hachures in Fig. 5 is a constraint in parameter space which implies that only a low-viscosity contrast ($< 10^2$) and a viscosity of upper mantle under spreading centres of $\mu_2 = 10^{18 \pm 1} \text{ P}$ are compatible with the theory and the observed length- and timescales of magmatic activity on mid-ocean ridges. The centre of intersecting hachures in Fig. 5 lies at $\sim \mu_2 = 10^{18} \text{ P}$, ($\mu_2/\mu_1 = 10$, $c = 3\%$).

These estimates are consistent with some geophysical, petrological and experimental observations. From gravity and models of the median valley on the MAR, Collette *et al.*³⁵ have estimated a value of $10^{19.5} \text{ P}$ for the viscosity of upper mantle beneath this plate boundary. Based on petrological studies of Alpine peridotites, Dick and Sinton³⁶ concluded that dissolution creep may be the dominant mechanism for both deformation and compositional layering in a partially molten upper mantle under spreading centres. Recent hot-pressing experiments by Cooper and Kohlstedt³⁷ on olivine-basalt assemblages to simulate dissolution creep in a peridotitic upper mantle indicate that the viscosity contrast between melt-free and partially molten mantle would be > 1 and < 10 . This suggests that the choice of observed length- and timescales of magmatic activity on mid-ocean ridges and the proposed theory are appropriate, and if tighter constraints on these scales and on μ_2 , μ_1 , c or $\Delta\rho$ could be achieved, a more refined theory (tested carefully by laboratory or numerical experiments) may lead to a clearer understanding of upwelling under spreading centres.

From this study we conclude that any spreading boundary in the world's ocean basins would exhibit an even and predictable segmentation. Evaluation on the proposed 30–85-km length scale of other existing spreading centre data, as well as additional high-resolution surveys of seafloor spreading boundaries are required to determine whether spreading-centre spacing is dependent only on spreading rate (indicating a uniform process of upwelling beneath ridges on a global scale) or whether a consistent spreading-centre spacing and consequent uniformity of mantle upwelling should be considered on the regional scale of individual plate boundaries, mantle plumes or other heterogeneities of the Earth's mantle.

We thank H. J. B. Dick, J. A. Karson, R. C. Kerr and G. M. Purdy for constructive criticism. This work was supported by the Office of Naval Research, Contract N00014-82-C-0019 and a grant from the Center for Analysis of Marine Systems (CAMS) of the Woods Hole Oceanographic Institution with funds from the Exxon Education Foundation. Pam Foster and Maureen Carragher processed the manuscript. Woods Hole Oceanographic Institution Contribution 5974.

Received 18 March; accepted 17 July 1985.

- Schouten, H. & Klitgord, K. D. *Earth planet. Sci. Lett.* **59**, 255–266 (1982).
- Schouten, H. & Klitgord, K. D. *Nature* **303**, 549–550 (1983).
- Schouten, H. & White, R. S. *Geology* **8**, 175–179 (1980).
- Detrick, R. S. & Purdy, G. M. *J. geophys. Res.* **85**, 3759–3777 (1980).
- Detrick, R. S., Cormier, M. H., Prince, R. A., Forsyth, D. W. & Ambos, E. L. *J. geophys. Res.* **10599–10612** (1982).
- Sinha, M. C. *EOS* **64**, 314 (1983) (abstr.).
- Mutter, J. C. & Detrick, R. S. *Geology* **12**, 534–537 (1984).
- Macdonald, K. C. & Fox, P. J. *Nature* **302**, 55–58 (1983).
- Macdonald, K. C., Sempere, J.-C. & Fox, P. J. *J. geophys. Res.* **11**, 6049–6069 (1984).
- Whitehead, J. A., Dick, H. J. B. & Schouten, H. *Nature* **312**, 146–148 (1984).
- Howard, L. N. *Proc. 11th int. Congr. Appl. Mech.*, Munich, 1109–1115 (Springer, Berlin, 1966).
- Lonsdale, P. *EOS* **63**, 1108 (1982).
- Lonsdale, P. *J. geophys. Res.* **88**, 9393–9406 (1983).
- Lonsdale, P. *Bull. geol. Soc. Am.* **96**, 313–327 (1985).
- Ramberg, I. B. & van Andel, T. H. *Bull. geol. Soc. Am.* **88**, 577–586 (1977).
- Rona, P. A., Harbison, R. N., Bussiner, B. G., Scott, R. B. & Natwalk, A. J. *Bull. geol. Soc. Am.* **87**, 661–674 (1976).
- Rona, P. A. & Gray, D. F. *Bull. geol. Soc. Am.* **91**, 485–496 (1980).

- Fox, P. J. & Gallo, D. G. *Tectonophysics* **104**, 205–242 (1984).
- Karson, J. & Dick, H. J. B. *Mar. geophys. Res.* **6**, 51–98 (1983).
- Ballard, R. D., van Andel, T. H. & Holcomb, R. T. *J. geophys. Res.* **87**, 1149–1161 (1982).
- Francheteau, J. & Ballard, R. D. *Earth planet. Sci. Lett.* **64**, 93–116 (1983).
- Lichtman, G. S. & Eissen, J. P. *Geology* **11**, 592–595 (1983).
- Crane, K. *Earth planet. Sci. Lett.* **72**, 405–414 (1985).
- Whitehead, J. A. & Luther, D. S. *J. geophys. Res.* **80**, 705–717 (1975).
- Marsh, B. D. *J. Geol.* **87**, 687–713 (1979).
- Bryan, W. B., Thompson, G., Frey, F. A. & Dickey, G. S. *J. geophys. Res.* **81**, 4285–4304 (1976).
- Turcotte, D. L. & Schubert, G. *Geodynamics Applications of Continuum Physics to Geological Problems* (Wiley, New York, 1982).
- Parker, R. L. & Oldenburg, D. W. *Nature Phys. Sci.* **242**, 137–139 (1973).
- Klitgord, K. D. & Mammertick, J. *J. geophys. Res.* **87**, 6725–6750 (1982).
- Searle, R. C. *Geology* **11**, 607–610 (1983).
- Palm, E. *Adv. appl. Mech.* **18**, 77–121 (1978).
- Busse, F. M. *Adv. appl. Mech.* **18**, 77–121 (1978).
- Scott, D. R. & Stevenson, D. J. *Geophys. Res. Lett.* **11**, 1161–1164 (1984).
- Ballard, R. D. & van Andel, T. H. *Bull. geol. Soc. Am.* **88**, 507–530 (1977).
- Collette, B. J., Verhoef, J. & Mulder, A. F. J. *J. Geophys.* **47**, 91–98 (1980).
- Dick, H. J. B. & Sinton, J. M. *J. Geol.* **87**, 403–416 (1979).
- Cooper, R. F. & Kohlstedt, D. L. *Tectonophysics* **107**, 207–233 (1984).
- Minster, J. B. & Jordan, T. H. *J. geophys. Res.* **83**, 5331–5354 (1978).

Insertion of DNA sequences into the human chromosomal β -globin locus by homologous recombination

Oliver Smithies*, Ronald G. Gregg*, Sallie S. Boggs†, Michael A. Koralewski* & Raju S. Kucherlapati‡

* Laboratory of Genetics, University of Wisconsin, Madison, Wisconsin 53706, USA

† School of Medicine, University of Pittsburgh, Pittsburgh, Pennsylvania 15261, USA

‡ Center for Genetics, University of Illinois College of Medicine, Chicago, Illinois 60612, USA

A 'rescuable' plasmid containing globin gene sequences allowing recombination with homologous chromosomal sequences has enabled us to produce, score and clone mammalian cells with the plasmid integrated into the human β -globin locus. The planned modification was achieved in about one per thousand transformed cells whether or not the target gene was expressed.

VARIOUS methods are available for introducing exogenous DNA into mammalian cells¹⁻⁵, where it can be stably incorporated into the genome^{6,7}. The sites of incorporation have not, however, been under control. Results that could be interpreted as signifying recombination between incoming DNA and a homologous chromosomal gene have been reported by Goodenow *et al.*⁸, who transfected mouse L cells with truncated *H-2* genes and observed the expression of full-length *H-2* antigens. They did not test this interpretation. Yoshie *et al.*⁹ have made similar observations in experiments with truncated *HLA* genes; limited DNA sequence analysis showed that homologous recombination with known *HLA* genes could not account for their data.

The series of experiments reported here shows that homologous recombination between exogenous DNA and a chromosomal gene does occur in cultured mammalian cells and can allow the stable insertion of defined DNA sequences into the human β -globin locus in a predictable fashion. Although the frequency of success ($\sim 10^{-3}$ of the transformed cells) is at present modest, the data show unequivocally that the planned modification of a specific human gene is possible *in vivo*.

We chose the human β -globin locus as a target for gene modification for three reasons: (1) the desirability of eventually being able to correct harmful mutations at this locus in patients with thalassaemia and sickle-cell anaemia; (2) mammalian bone marrow stem cells are likely to become available within the foreseeable future in quantities and purity sufficient to allow tests of the targeted modification with stem cells; and (3) the availability of cloned and sequenced DNA covering much of the locus^{10,11}.

The rationale used to direct the incoming DNA to the β -globin locus is derived from experiments in yeast by Hinnen *et al.*¹² showing that recombination between incoming DNA sequences and homologous sequences within the genome can direct the exogenous DNA to the desired locus. A refinement, based on work in yeast by Orr-Weaver *et al.*¹³, is to cut the incoming DNA within the region of homology using a suitable restriction enzyme. In a previous study¹⁴, we showed that the provision of such recombinogenic ends greatly increases homologous recombination between plasmids in mammalian somatic cells. This suggests that cutting within the region of homology may also significantly increase the frequency of directed gene modification in mammalian cells.

Detecting target gene modification

Figure 1 shows the planned modification to the human β -globin locus, that is, the integration by homologous recombination of a single copy of a test plasmid into the locus at a defined position. To achieve this modification, we constructed a special test plasmid, p $\Delta\beta$ 117, that could also be used in a sensitive screening

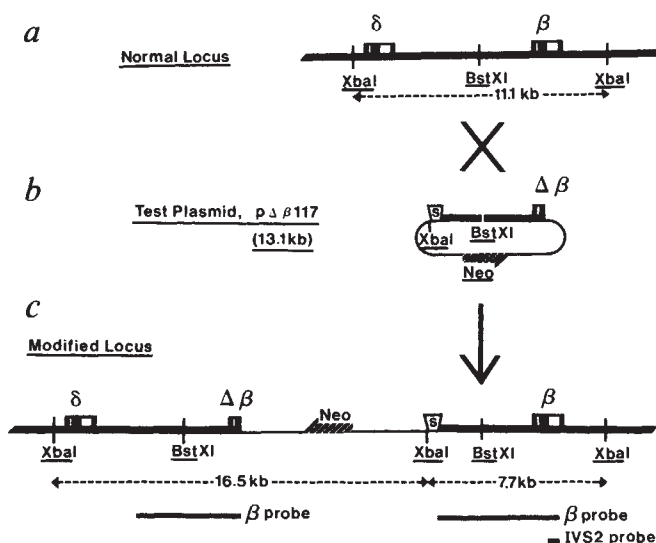


Fig. 1 Planned modification of the human β -globin locus. *a*, Map of normal human DNA encompassing the adult δ - and β -globin genes, with their coding regions and introns indicated by solid and open bars. The transcriptional sense of these genes is from left to right. All *Xba*I sites are shown. Only one of two *Bst*XI sites in the region is indicated; the other, omitted for clarity, is in the third exon of the β -globin gene. *b*, The 13.1-kb test plasmid, p $\Delta\beta$ 117, used as a vehicle to achieve and detect the targeted modification. Human globin locus DNA sequences (4.6 kb) in the test plasmid are indicated by a heavy line that is aligned appropriately with *a*. (The sequences extend from the first *Eco*RI site upstream of the β -globin gene to the *Bam*HI site at the 3' end of the second exon.) The 5' part ($\Delta\beta$) of the β -globin gene within these sequences includes the first exon and most of the second exon (shown by solid bars) and the first intervening sequence (open bar) but does not include the second intervening sequence, IVS-2. The light line shows sequences largely derived from the plasmid pSV2Neo¹⁶ from which p $\Delta\beta$ 117 was constructed. The *neo* gene is indicated by a crosshatched arrow (not to scale) that also shows its transcriptional orientation. The boxed letter S (not to scale) represents the *supF* gene, which encodes a tyrosine tRNA able to suppress some amber mutations in *E. coli*. The *supF* gene was taken from the plasmid p π VX of Seed¹⁷ together with a polylinker that provided the adjacent *Xba*I site. The naturally occurring *Bst*XI site within the globin sequences of p $\Delta\beta$ 117 is indicated; this is the only *Bst*XI site within the test plasmid. It was cut as indicated before introducing the plasmid into recipient cells. *c*, Map of the β -globin locus as it should appear after the planned modification by homologous recombination. Horizontal bars at the bottom of *c* show regions in the modified β -globin locus to which the two probes used in the assay will hybridize.

Table 1 Fragment rescue from cells not expressing (EJ) or expressing (Hu 11) the β -globin gene

Cells	G418-resistant colonies	Total genomic DNA assayed (μ g)	Total phage*	Phage growing on C-1a	Phage hybridizing to β probe	Phage hybridizing to β and IVS-2 probes†
EJ carcinoma (no G418)	[405]‡	100	6.3×10^8	2,520	125	1
EJ carcinoma (no G418)	[1,530]‡	50	3.4×10^8	820	24	1
Hu 11 hybrid (+G418)	4,400	~20 (5.5–8.5-kb fraction)	2.4×10^8	1,420	145	2
Hu 11 hybrid (+G418)	4,400	~20 (8.5–16.5-kb fraction)	1.4×10^8	1,760	60	0

Cells were transformed with p Δ 117 cut with *Bst*XI and grown in the absence or presence of G418, as indicated below. DNA was prepared, digested with *Xba*I and assayed for fragment rescue. Nine dishes of human EJ bladder carcinoma cells (5×10^5 per dish) were treated with calcium phosphate precipitates of carrier-free p Δ 117 DNA, linearized by cleavage with *Bst*XI, at either 1 μ g (first line) or 8 μ g (second line) per dish. After 3–4 days of culture in medium without G418, the cells were collected for DNA preparation and the DNA was digested with *Xba*I. 10^8 Hu 11 hybrid cells (see text), grown in HAT medium, were suspended in 1 ml 140 mM NaCl, 25 mM HEPES, 0.75 mM Na_2HPO_4 , pH 7.15, in the presence of 25 μ g ml $^{-1}$ *Bst*XI-digested p Δ 117 DNA. Electroporation^{4,24} was carried out in an ice-cooled 0.5 ml treatment vessel with 1 cm 2 platinum end-electrodes. A solid state electronic switch was used to discharge a bank of condensers (effective capacitance 14 μ F charged to 1,000 V) through the treatment vessel. HAT medium was added 10 min after the shock and the cell suspension was dispensed into four culture flasks, except for a small amount used for estimating transformation frequencies. Two days later an equal volume of HAT medium containing G418 was added to all cultures to give a final concentration of 400 μ g G418 ml $^{-1}$. Subsequent medium changes every three days used HAT medium plus 400 μ g G418 ml $^{-1}$. DNA was prepared from the cells grown in each of the four culture flasks (estimated to contain 1,100 G418-resistant colonies per flask), and from these samples a pool of DNA was prepared equivalent to 4,400 colonies. After *Xba*I digestion the pooled DNA was separated by agarose electrophoresis into two fractions: 5.5–8.5 kb and 8.5–16.5 kb. The *Xba*I-digested DNA in all the experiments was ligated to DNA from the double-amber mutant bacteriophage Charon 3A Δ Xba, made from Charon 3A Δ Lac²⁵ by adding an *Xba*I cloning site and deleting 1.8 kb (net) of DNA to increase its cloning capacity to ~11 kb. The ligated DNA was packaged²⁶ into viable phages using sonic extracts and freeze-thaw lysates prepared from *recA* $^-$ strains of *E. coli*. The total number of phages was assayed on *E. coli* strain K802, which suppresses the amber mutations. Phages that rescued the *supF* gene from the genomic DNA or lost their amber mutations during the packaging (see text) were assayed on late logarithmic cells of *E. coli* strain C-1a, which cannot suppress amber mutations. Phage plaques on strain C-1a were blotted twice with nitrocellulose and hybridized to two probes (see Fig. 1): a β probe which will hybridize to the human globin sequences in the test plasmid p Δ 117, and an IVS-2 probe which will hybridize only to phages containing the 3' part of the β -globin gene that was not in the incoming test plasmid. (The β probe is a 5,548-bp *Eco*RI fragment that includes the 5' part of the human β -globin gene; the IVS-2 probe is a 917-base pair *Bam*HI/*Eco*RI fragment that spans the β -globin IVS-2 sequence.)

* This total, obtained from titration on *E. coli* K802, includes recombinant and non-recombinant phages, because the Charon 3A Δ Xba vector phage will grow on K802 with or without an insert.

† These doubly hybridizing phages were examined at the DNA level by electrophoresis and were shown to contain the predicted 7.7-kb *Xba*I fragment.

‡ The cells used for DNA preparation were not exposed to G418. These estimates of G418-resistant cells are calculated from small parallel samples.

assay to detect and measure the frequency of the targetted modification¹⁵. Because it is a screening method, the assay allows a target gene to be chosen for which no direct selection method is available. It depends on rescuing DNA sequences corresponding to the input DNA from the genome of the recipient cells, together with genomic sequences next to this DNA, and then determining whether the adjacent sequences correspond to the expected region of the target locus.

The test plasmid p Δ 117 incorporates four important elements. The first is 4.6 kilobases (kb) of DNA from the human β -globin locus, including the 5' part of the adult human β -globin gene but lacking the second intervening sequence (IVS-2) and the remaining 3' portion of the gene. This DNA provides sequences in the incoming test plasmid that can recombine with homologous target sequences in the β -globin locus of the recipient cell to accomplish the planned modification. Within the 4.6-kb element is a single *Bst*XI site used to cut the incoming plasmid in the region homologous with the target.

The second element is sequences from the plasmid pSV2Neo of Southern and Berg¹⁶ that specify the *neo* gene and associated transcriptional signals. These sequences can confer resistance to G418 (an antibiotic related to neomycin) on mammalian cells. In some experiments, we used G418 to select recipient cells that had incorporated the test plasmid in a configuration in which the *neo* gene was functional.

The third element in p Δ 117 is the *supF* gene, which encodes a tyrosine transfer RNA that suppresses several amber mutations in *Escherichia coli* and its bacteriophages. When this suppressor gene is inserted into a mutant bacteriophage, carrying amber mutations preventing the phage from growing, the amber mutations will be suppressed and the phage will grow. We used a

double-amber mutant bacteriophage cloning vector (Charon 3A Δ Xba¹⁵) in a variation of a scheme described by Seed¹⁷ to rescue *supF*-containing sequences from the recipient mammalian cell genome. Only those phages that have rescued the *supF* gene will, in principle, grow on a wild-type strain of *E. coli* (C-1a) that lacks amber suppressors. (In practice, phages also grow that have lost both their amber mutations by recombination during packaging of the phage DNA¹⁸, but these phages do not contain any globin sequences and so do not compromise the assay).

The fourth important element in p Δ 117 is a site for the restriction endonuclease *Xba*I located upstream of the *supF* gene relative to the human β -globin sequences. In conjunction with a second *Xba*I site downstream of the adult β -globin gene in the normal human genome, this *Xba*I site generates a novel 7.7-kb *Xba*I fragment of DNA when homologous recombination has accomplished the planned modification (see Fig. 1). The 7.7-kb *Xba*I fragment is unique among the *supF*-containing genomic fragments rescued during the assay because it is the only one that contains both the *supF* gene and the IVS-2 part of the β -globin gene.

Thus, the overall procedure consists of the following stages: (1) introduction of the test plasmid DNA into the recipient cells; (2) in most experiments, selection with G418 of those cells that have incorporated the test plasmid into their genomes; (3) preparation of DNA from pools of these transformed cells; (4) digestion of DNA from the cell pools with the restriction enzyme *Xba*I; (5) ligation of the resulting genomic DNA fragments to DNA from the double-amber mutant vector phage (Charon 3A Δ Xba), packaging of the ligated DNA into viable phages and plating them on *E. coli* C-1a which lacks any suppressor genes;

Table 2 Fragment rescue from pools of transformed Hu 11 hybrid cells

Pool description	Number of colonies	Total genomic DNA assayed (μ g)	Total phage*	Phage growing on C-1a	Phage hybridizing to β probe	Phage hybridizing to β and IVS-2 probes
Flask 1	1,100†	40	1.3×10^9	2,650	381	1
Flask 2	1,100†	20	2.6×10^8	800	60	1
Flask 3	1,100†	20	2.7×10^8	570	61	2
Flask 4	1,100†	40	9.4×10^8	1,630	265	0
Flask 5	300‡	20	2.2×10^8	150	25	3
Flask 6	100‡	~40	2.8×10^8	440	41	0
Pool 1	20‡	40	7.5×10^7	115	36	8¶
Pool 2	20‡	20	3.3×10^7	53	22	0
Pool 3	20‡	20	2.2×10^7	25	11	0

Hu 11 hybrid cells were transformed by electroporation to G418 resistance as in Table 1, except as noted below, and were cultured in flasks or as pools of individual colonies as indicated. DNA was isolated, digested with *Xba*I and assayed as described in Table 1.

* Total includes recombinant and non-recombinant phages.

† Estimated from a small sample.

‡ Pooled from individual G418-resistant colonies obtained following electroporation in the presence of $12.5 \mu\text{g ml}^{-1}$ p Δ 117 cut with *Bst*XI.

|| Cells from a frozen sample of flask 5 were thawed, recloned, amplified and pools of 20 individual colonies were prepared; as a consequence of this procedure the colonies in these pools are no longer independent.

¶ Representative phages were examined at the DNA level by electrophoresis, and were shown to contain the predicted 7.7-kb *Xba*I fragment. The *Xba*I fragment from one doubly hybridizing phage was subcloned and shown, by further digests, to have the expected map, including the *supF* gene.

(6) scoring the resulting phages for hybridization to two probes (see Fig. 1). The first probe (the β -probe) identifies phages that have rescued any human globin sequences derived from the incoming test plasmid. The second probe (the IVS-2 probe) is made from the IVS-2 portion of the human β -globin gene that was absent from the incoming test plasmid but will be next to plasmid-derived globin sequences after the planned modification. Doubly hybridizing phages detect and score the occurrence of the modification. They can be purified and further checked by electrophoresis to ensure that they contain the predicted 7.7-kb *Xba*I fragment.

Preliminary work

In initial experiments¹⁵ we had transformed human EJ bladder carcinoma cells to G418 resistance with various test plasmids of the same general design as and including p Δ 117, but no successful modification of the native β -globin gene was detected in ~3,000 independent G418-resistant colonies sampled. This result was surprising because we were successful when using an incoming plasmid to correct a defective chromosomal plasmid that we had introduced into the genome as an artificial target¹⁵ (see also similar experiments by Smith and Berg¹⁹ and Lin *et al.*^{20,21} and related experiments with viral sequences by Shaul *et al.*²²). A possible explanation for this initial failure is that the combination of G418 selection and bladder carcinoma cells that do not express their β -globin genes is inappropriate. Specifically, any carcinoma cells in which the desired modification has been achieved might not survive G418 selection if the location of the inserted DNA within an inactive chromosomal locus prevents expression of the *neo* gene. Two sets of tests were carried out to avoid this problem. In the first, G418 selection was omitted. In the second, we used recipient cells in which a human adult β -globin gene was being expressed.

Tests without G418 selection

Human EJ bladder carcinoma cells not expressing their β -globin genes were used as recipients for incoming plasmid DNA in two experiments omitting G418 selection. Details of these experiments and their results are presented in Table 1. One phage from each experiment hybridized to both probes. Each was plaque-purified and its DNA was shown to contain the expected 7.7-kb *Xba*I fragment. Additional digests with *Eco*RI and *Bam*HI confirmed that the *Xba*I fragment had the map expected for a complete β -globin gene¹¹ including IVS-2. These experiments are therefore strong evidence that the locus has

been modified in the predicted fashion in these cells not expressing the β -globin gene, and that the modification can be detected when G418 selection is omitted. Although the experiments were based on the supposition that the *neo* gene cannot be expressed when inserted into an inactive β -globin locus, they do not prove that the supposition is correct.

Cells expressing a human β -globin gene

The second set of tests required a clonable recipient cell expressing the human adult β -globin gene. A cell line described by Zavodny *et al.*²³ meets these criteria. It was made by fusing hypoxanthine phosphoribosyltransferase-deficient tetraploid mouse erythroleukaemia cells to diploid human fibroblasts bearing an X;11 translocation chromosome and selecting for growth in medium containing hypoxanthine, aminopterin and thymidine (HAT medium). The resulting hybrid cell line, which we term the Hu 11 hybrid, expresses the adult β -globin gene on human chromosome 11 at a measurable level in the absence of induction, and at a high level after induction with dimethyl sulphoxide²³.

Hu 11 hybrid cells were transformed to G418 resistance by p Δ 117 cut with *Bst*XI using electroporation⁴ as described in Table 1. A pool of DNA representing 4,400 G418-resistant colonies of transformed cells was prepared and digested with *Xba*I. Agarose gel electrophoresis was used to prepare two size fractions from this DNA. One fraction covered the size range 5.5–8.5 kb and the other the range 8.5–16.5 kb. Using this size-fractionated DNA served two functions: it reduced the total amount of DNA to be packaged and it eliminated a conceivable artefact, namely that the diagnostic 7.7-kb *Xba*I fragment might be formed from the 11-kb *Xba*I fragment that includes the β -globin locus in unmodified recipient cells (see Fig. 1) by homologous recombination during the phage packaging²⁰ or subsequent steps of the assay.

The results obtained with the two fractions are shown in Table 1. Two doubly hybridizing phage were detected with the 5.5–8.5-kb fraction but none was found in the 8.5–16.5-kb fraction. Because the doubly hybridizing phages were obtained with <8.5 kb DNA, this result excludes the possibility that they were formed from the unmodified 11-kb *Xba*I fragment during the assay. The doubly hybridizing phages were plaque-purified and shown to contain the predicted 7.7-kb *Xba*I fragment. These experiments thus provide strong evidence that the planned modification has been achieved in Hu 11 hybrid cells expressing the β -globin locus and that G418 selection does not kill the

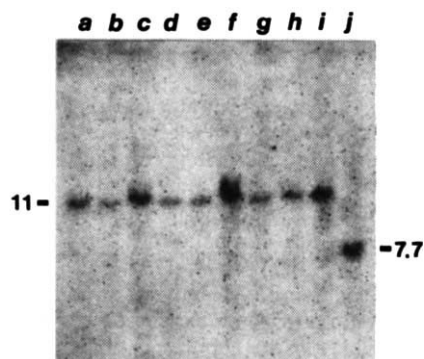


Fig. 2 Nitrocellulose blot of *Xba*I digests (lanes a-j) of DNA from 10 G418-resistant colonies (A-J) from pool 1 (see text and Table 2) hybridized to the IVS-2 probe specific for the human β -globin gene. Measured fragment sizes are shown in kb.

resulting cells. Accordingly, we used the assay in conjunction with a sib-selection procedure to identify and isolate from the G418-resistant cells a clone containing the modified β -globin locus.

Clone isolation

The pooled DNA studied in the preceding section was made from four flasks of cells each containing an estimated 1,100 G418-resistant colonies. In similar experiments we prepared two flasks containing pools of 300 or 100 isolated individual colonies, respectively (see Table 2 for data obtained with these six flasks). Three of the 1,100-colony flasks and the 300-colony flask scored positively, whereas the other two flasks were negative. Thus, the presumptive modification is being accomplished about once per 300 to once per 1,100 cells transformed to G418-resistance.

Because the 300-colony flask scored positively, we thawed a frozen sample of cells from this flask and recloned them. The resulting clones were expanded and samples were taken to create several 20-colony pools. (Note that because these colonies are subclones from one flask, they are no longer independent.) DNA from three of the 20-colony pools was tested; pool 1 scored positively whereas pools 2 and 3 did not (Table 2). Accordingly, we prepared and tested directly DNA samples from each of the 20 individual colonies constituting pool 1.

Direct test of modification

Figure 1 shows how the planned addition of the 13.1-kb test plasmid to the human β -globin locus will alter the restriction map of the locus. The most obvious predictions are that the diagnostic 7.7-kb *Xba*I fragment should be detectable by the β -globin IVS-2 probe in Southern blots of DNA from the modified cells, and that this fragment should replace an 11-kb *Xba*I fragment present in the DNA of unmodified cells.

Figure 2 shows a Southern blot of *Xba*I digests of DNA from 10 of the 20 colonies comprising pool 1. In one of the colonies (J) the predicted change is seen; a 7.7-kb *Xba*I fragment has replaced the usual 11-kb fragment from the unmodified locus. (The same result was also seen in another colony, P, in the remaining 10 colonies of pool 1.) As no residual 11-kb fragment is present, the modified colony probably has only one copy of the β -globin locus. (Karyotyping of comparably treated Hu 11 hybrid cells showed only one chromosome with the banding pattern of human chromosome 11.) Thus, the fragment rescue assay for the planned modification is validated in that it led us to two not necessarily independent colonies in which the modification is demonstrable by direct tests.

Fidelity of modification

To determine whether the modification is the sought-for simple one-copy insertion of p $\Delta\beta$ 117, digests of genomic DNA from one of the modified colonies (P) were hybridized to several

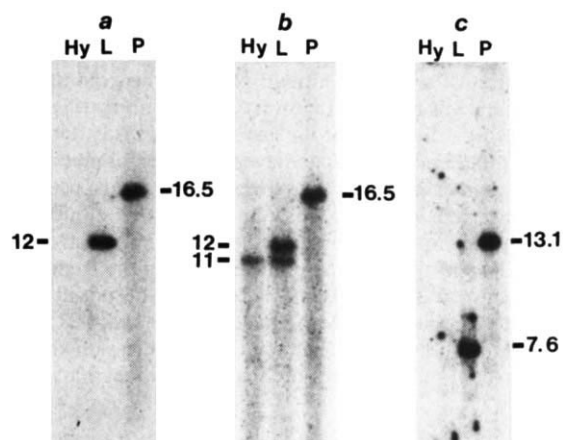


Fig. 3 Nitrocellulose blots of digests of DNA from two of the G418-resistant colonies from pool 1 (see text and Table 2) and from untreated Hu 11 hybrid cells. The blots were: a, *Xba*I digests hybridized to a probe specific for pSV2Neo-derived sequences; b, *Xba*I digests hybridized to probes specific for the IVS-2 of the human δ -globin gene and for pSV2Neo sequences; c, *Bst*XI digests hybridized to a probe specific for pSV2Neo sequences. Samples are: Hy, untreated Hu 11 hybrid cells; L, the G418-resistant colony L; P, the G418-resistant colony P. Measured fragment sizes are shown in kb.

probes (Fig. 3). As controls, DNA was included from untreated Hu 11 hybrid cells and from one of the other colonies (L) in which p $\Delta\beta$ 117 was incorporated at a non-globin location.

Figure 3a shows *Xba*I digests of the samples hybridized to a probe (pSV2Neo) that will detect most of the non-globin parts of inserted p $\Delta\beta$ 117 sequences. The modified colony P gives only one hybridizing band at 16.5 kb. This shows that the map is altered as expected (see Fig. 1) from the insertion of only a single copy of p $\Delta\beta$ 117 between the δ - and β -globin genes, and it also shows that no plasmid sequences were incorporated into other (illegitimate) sites in the genome of the modified colony. Colony L also gives a single hybridizing band indicating the incorporation of p $\Delta\beta$ 117 sequences into the recipient cell genome at only a single location, but the size of the fragment (12 kb) demonstrates that the incorporation was not at the targeted site. The untreated Hu 11 hybrid cells show no plasmid sequences.

Figure 3b, obtained by hybridizing a probe specific for the human δ -globin gene to the blot used for a (already hybridized to pSV2Neo) confirms and extends these conclusions. Colony P still shows only the single 16.5-kb band that contains both plasmid sequences and the δ -globin gene. In contrast, colony L now shows two bands, because the plasmid sequences and the δ -globin gene are on different *Xba*I fragments (12 and 11 kb, respectively). The untreated Hu 11 hybrid cells show only the 11-kb fragment from the unmodified locus.

Because *Bst*XI was used to create recombinogenic ends on the incoming plasmid, a *Bst*XI digest of DNA from the modified colony is particularly informative. Accurate insertion of a single copy of the test plasmid without any net degradation should result in two *Bst*XI sites (see Fig. 1) separated by the length of the plasmid (13.1 kb). Figure 3c shows that such a 13.1-kb fragment, detected by the plasmid-specific pSV2Neo probe, is indeed present in *Bst*XI digests of DNA from the modified colony P. Colony L, having the plasmid inserted at an illegitimate site, gives a different and unpredictable *Bst*XI fragment at 7.6 kb.

We conclude that the modification observed in the Hu 11 hybrid cells was the sought-for insertion of a single copy of the test plasmid into the β -globin locus. As the modified colonies are G418 resistant, our results also establish that the *neo* gene inserted into an active β -globin locus can be expressed even though its transcriptional orientation is opposite from that of the β -globin gene.

Conclusions

The experiments reported here establish that the planned modification of a specific human gene can be accomplished in mammalian cells by homologous recombination without detectably affecting other parts of the genome. Our assay data indicate that the modification was achieved whether or not the target gene was active. When it was active we were able to use selection of the inserted sequence to help clone the modified cells. Although many other problems remain to be solved, these results suggest that specific planned modifications of the β -globin locus in bone marrow stem cells may eventually be possible for the treatment of patients with haemoglobinopathies such as thalassaemia and sickle-cell anaemia. Regardless of whether this long-term goal can be made practical, it is evident that extensions

of the principles we have used here to other genes in complex eukaryotes should permit new approaches to many problems in molecular genetics.

This work was supported by grants to O.S. (AM-20120, GM-20069) and to R.S.K. (GM-33943, March of Dimes I-806). We thank Deb Endean and Paula Henthorn for help with p Δ B117 and Charon 3A Δ Xba, Natalie Borenstein for tissue culture, Anne Hanneken and Charlie Blystad for assays, Brian Morse and Stacy Ehrlich for DNA preparations, Lisa Becker for freeze-thaw lysates, Art Skoultschi for Hu 11 hybrid cells, Huntington Potter and Bill Sugden for advice on electroporation and Karen Lyons for suggestions concerning our early failures. This is paper number 2834 from the Laboratory of Genetics, University of Wisconsin-Madison.

Received 1 July; accepted 30 July 1985.

- Wigler, M. *et al.* *Cell* **11**, 223-232 (1977).
- Capecchi, M. R. *Cell* **22**, 479-488 (1980).
- Gordon, J. W., Scangos, G. A., Plotkin, D. J., Barbosa, J. A. & Ruddle, F. H. *Proc. natn. Acad. Sci. U.S.A.* **77**, 7380-7384 (1980).
- Neumann, E., Schaefer-Ridder, M., Wang, Y. & Hofschneider, P. H. *EMBO J.* **1**, 841-845 (1982).
- Miller, A. D., Jolly, D. J., Friedmann, T. & Verma, I. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4709-4713 (1983).
- Palmiter, R. D. *et al.* *Nature* **300**, 611-615 (1982).
- Chada, K. *et al.* *Nature* **314**, 377-380 (1985).
- Goodenow, R. S. *et al.* *Nature* **301**, 388-394 (1983).
- Yoshie, O. *et al.* *Proc. natn. Acad. Sci. U.S.A.* **81**, 649-653 (1984).
- Fritsch, E. F., Lawn, R. M. & Maniatis, T. *Cell* **19**, 959-972 (1980).
- Poncz, M., Schwartz, E., Ballantine, M. & Surrey, S. *J. biol. Chem.* **258**, 11599-11609 (1983).
- Hinnen, A., Hicks, J. B. & Fink, G. R. *Proc. natn. Acad. Sci. U.S.A.* **75**, 1929-1933 (1978).
- Orr-Weaver, T. L., Szostak, J. W. & Rothstein, R. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 6354-6358 (1981).

- Kucherlapati, R. S., Eves, E. M., Song, K.-Y., Morse, B. S. & Smithies, O. *Proc. natn. Acad. Sci. U.S.A.* **81**, 3153-3157 (1984).
- Smithies, O., Koralewski, M. A., Song, K.-Y. & Kucherlapati, R. S. *Cold Spring Harb. Symp. quant. Biol.* **49**, 161-170 (1984).
- Southern, P. J. & Berg, P. *J. molec. appl. Genet.* **1**, 327-341 (1982).
- Seed, B. *Nucleic Acids Res.* **11**, 2427-2445 (1983).
- Kataoka, T., Takeda, S.-I. & Honjo, T. *Proc. natn. Acad. Sci. U.S.A.* **80**, 2666-2670 (1983).
- Smith, A. J. H. & Berg, P. *Cold Spring Harb. Symp. quant. Biol.* **81**, 171-181 (1984).
- Lin, F.-L., Sperle, K. K. & Sternberg, N. *Cold Spring Harb. Symp. quant. Biol.* **81**, 139-149 (1984).
- Lin, F.-L., Sperle, K. & Sternberg, N. *Proc. natn. Acad. Sci. U.S.A.* **82**, 1391-1395 (1985).
- Shaul, Y., Laub, O., Walker, M. D. & Rutter, W. J. *Proc. natn. Acad. Sci. U.S.A.* **82**, 3781-3784 (1985).
- Zavodny, P. J., Roginski, R. S. & Skoultschi, A. I. *Globin Gene Expression and Hematopoietic Differentiation* (eds Stamatoyannopoulos, G. & Nienhuis, A. W.) 53-62 (Liss, New York, 1983).
- Potter, H., Weir, L. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **81**, 7161-7165 (1984).
- Blattner, F. R. *et al.* *Science* **196**, 161-169 (1977).
- Hohn, B. *Meth. Enzym.* **68**, 299-309 (1979).

LETTERS TO NATURE

New upper bound on the flux of cosmic magnetic monopoles

A. D. Caplin, C. N. Guy, M. Hardiman*, J. G. Park† & J. C. Schouten

Blackett Laboratory, Imperial College, London SW7 2BZ, UK

Grand unification theories (GUTs) suggest that massive magnetic monopoles (mass $\sim 10^{16}$ GeV) were created in the very early stages of the formation of the Universe. The report, 3 years ago¹, of a candidate cosmic magnetic monopole therefore stimulated intense experimental and theoretical activity^{2,3}. These monopoles should have the magnetic charge h/e predicted by Dirac⁴, or perhaps small integer multiples of h/e , but zero electric charge. Surviving magnetic monopoles are thought to be moving so slowly, $v/c = 10^{-3}$ or less, that they interact exceedingly weakly with matter. For this reason, conventional particle detectors (whose mechanism is electronic ionization or excitation) are expected to be rather insensitive⁵, so that inductive detectors are particularly attractive. We describe here the results of 6 months observation with a large inductive detector. Some putative events have been seen, but none of them seem to have been induced by a monopole. Our total exposure (detector area-time product) is now about 230 times that of Cabrera's first experiment, so that it seems increasingly likely that his original candidate event was spurious.

A Dirac monopole traversing an isolated closed conducting loop induces a current, by Faraday's law of induction, that generates magnetic flux equal to h/e ($=4.1 \times 10^{-15}$ Wb), equivalent to $2\Phi_0$, where Φ_0 is the magnetic flux quantum associated with superconductivity. No other particle will induce a net current, and furthermore the size of the current is independent of the trajectory and speed of the monopole. In practice, the detector loop is best made superconducting, so that the

induced current is permanent and more easily measurable. Fluctuations in the ambient magnetic field would also cause flux changes in the detector loop, so that good magnetic screening is essential. A convenient and effective way to do this is to surround the detector loop with a superconducting shield; however, the magnetic coupling (mutual inductance) between the loop and the shield complicates matters: the larger the size of the loop relative to the shield, the greater the coupling and the more the magnitude of the induced current in the loop depends on the details of the monopole trajectory.

Cabrera's original detector was a 50-mm diameter loop within a 200-mm diameter superconducting shield, and his detection of a single candidate event after 6 months of observation was at odds with the astrophysical limit⁶ on the number flux of cosmic monopoles derived from the persistence of galactic magnetic fields. Cabrera's single candidate monopole, if genuine and if interpreted as a rate, implied a number flux some 6 orders of magnitude greater than this astrophysical upper bound.

Our inductive detector contains three independent superconducting detector loops within a single superconducting shield, 960 mm high and 234 mm in diameter (Fig. 1). The two horizontal loops each have a single turn of diameter 170 mm and a counterwound 10-turn coil of diameter 54 mm to reduce the magnetic coupling to the shield⁷. The 'window-frame' (WF) loop, in contrast, is deliberately strongly coupled to the shield, but it is sensitive to all monopole trajectories intersecting the shield other than those nearly on axis⁸. While this design maximizes the effective detector area, it does sacrifice the valuable identifying feature of a single well-defined signal magnitude from passage of a monopole. There is partial coincidence between the three loops: trajectories that intersect either the top (T) or bottom (B) loops have high probability, $\approx 95\%$, of giving a significant signal in the WF loop. The probability of coincidence between B and T loops is $\approx 7\%$. Only about 10% of the tracks giving a signal in the WF loop intercept either the B or T loops. Each of the detectors is fitted with a toroidal calibration coil that can be energized externally, so as to simulate a monopole. Our inductive detector is, we believe, one of the two largest now in operation, being comparable in size with that

* Present address: School of Mathematical and Physical Sciences, University of Sussex, Falmer, Brighton BN1 9QH, Sussex, UK.

† Deceased 22 June 1983.

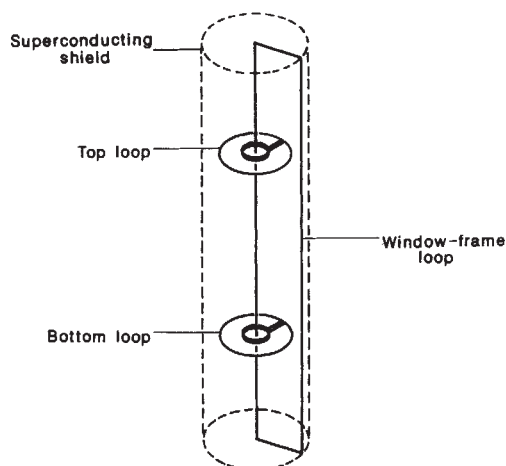


Fig. 1 Configuration of three independent loops in the Imperial College monopole detector. The superconducting shield is indicated by the broken line. For clarity, the SQUIDS (which are positioned near the top of the shield), calibration coils and other details have been omitted.

used by Incandela *et al.*⁹, although substantially larger detectors are being built elsewhere (refs 10, 11 and P. Chaudhari, personal communication).

The detector assembly sits in liquid helium in a large vapour-cooled cryostat that is surrounded by a set of five concentric mu-metal shields at room temperature. The ambient field in the detector region is thereby reduced to $\approx 10^{-9}$ T, and is held constant by the superconducting shield. The equipment is sited in the building basement, and is isolated as well as possible from vibration, RF (radiofrequency) interference, and so on. The currents in the three detectors are measured by commercial SQUID (Superconducting Quantum Interference Devices, SHE Corporation) sensors with bandwidths set to 10 Hz; their outputs, together with those from an accelerometer attached to the top of the cryostat, a magnetometer sensing the external magnetic field, an RF interference monitor, and a sensor on the helium pressure within the cryostat, are sampled 20 times per second by a computer-controlled data acquisition system. A simple step-detecting algorithm is applied to each of the detector channels; if it is triggered, full data on all channels for 300 s pre-trigger and 100 s post-trigger are dumped to disk for later analysis. Routinely, the computer plots out minimum and maximum values for each channel in every 100 s period, and it is supplemented by conventional chart recorders, operating at 1-Hz bandwidth, on the three detector outputs. Full details of the detector will be published elsewhere¹².

For an event to be considered as a candidate monopole, other possible causes, such as mechanical shock, must be absent. Also, because the transit time of even a very slow monopole would be fast compared with the rise-time of the electronics (65 ms for 0–90%), a candidate event must appear on the record as bandwidth-limited. For the B and T channels, the event magnitude should be close to $2\Phi_0$, and is almost certain to be accompanied by a simultaneous event in the WF channel. On the other hand, events in the WF channel can have any size between 0 and $1\Phi_0$ (Fig. 2), and there need not be coincidences with the B or T channels.

Our detector started taking data intermittently in May 1984, and has been running continuously since 28 September 1984, with just 2 or 3 h 'dead time' each week when the helium is refilled. As at 15 April 1985, total exposure is 4,850 h.

To date, some 90 'events' have been triggered. In most cases, the cause has been obvious immediately, for example: operator error, gross mechanical shock, or the helium level being allowed to fall below the top of the detector assembly. All three detectors appear to be fairly immune to magnetic and RF disturbance, but the WF loop is orders of magnitude more sensitive than the B and T loops to mechanical shock and to variations in the pressure above the helium bath; it is almost certain that the

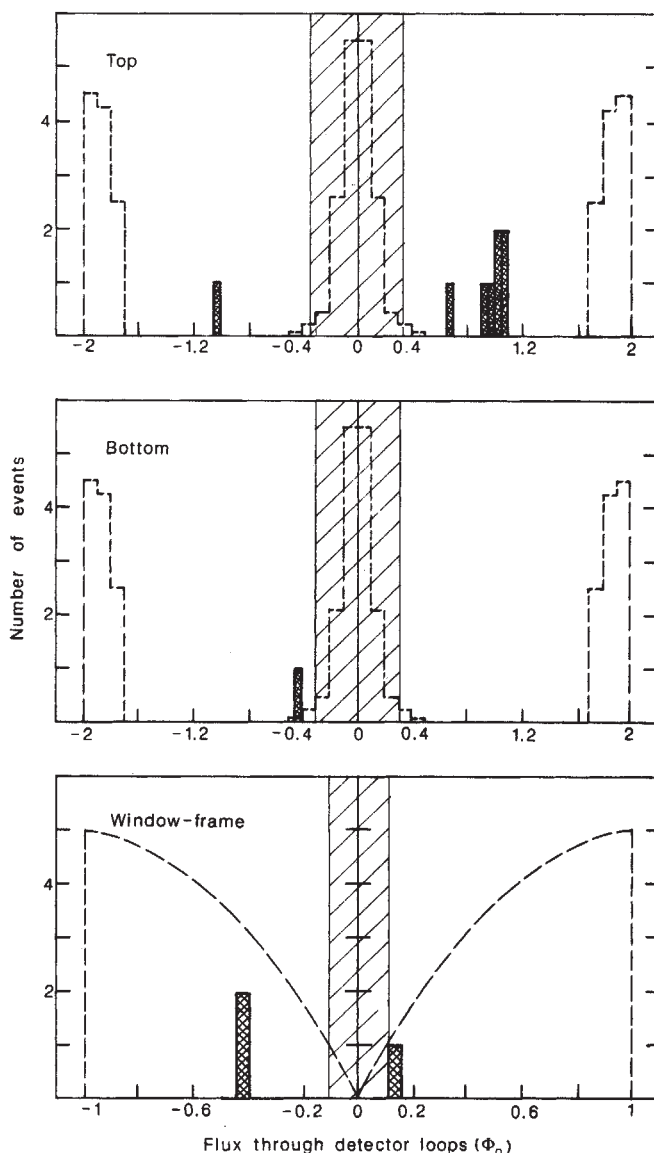


Fig. 2 Summary of the 12 'non-trivial' events recorded up to 15 April 1985. Their likely causes are discussed in the text, and none of them are serious candidate monopoles. The broken lines show the calculated distribution of event sizes for an isotropic influx of Dirac monopoles (the large peak near zero in the top and bottom loops arises from monopole trajectories that pass through the annular region between the loop and the shield); the hatched region near each origin represents events too small to trigger the data acquisition system.

pressure changes matter because of their indirect mechanical effect. The greater sensitivity of the WF loop arises because it is tightly coupled both mechanically and magnetically to the superconducting shield, and it therefore responds to any shield deformation, whereas the B and T loops are effectively decoupled from the shield.

External shocks of significant magnitude have been rare and readily identifiable. Much more troublesome have been the internal shocks generated by thermal expansion inside the cryostat as the helium level falls and the temperature profile changes. As low-temperature physicists know well, this expansion can be spasmodic and acoustically noisy. The accelerometer, which is mounted at the top of the cryostat, does pick up these internal shocks; however, because it is remote from the detector assembly, and sensitive to acceleration along a single axis only, it does not give a quantitative measure of the effect on the detector itself. These internal shocks have been another fairly common cause of events, particularly in the WF channel.

Elimination of the events with obvious causes leaves 12 non-

Table 1 Near-coincident shocks and events in the WF channel

Event no.	Event size (Φ_0)	Shock (10^{-3} g)	Delay between shock and event (s)	Ambient shock frequency (h^{-1})	Background
34	-0.43	10	143	1	Stable
35	-0.41	4	23	10	
74	+0.14	4	10	1	
42	+0.29	6	100	5	Drifting because of pressure changes
43	+0.18	5	-107	5	
44	+0.36	6	-65	5	
56	+0.22	10	103	<1	

The shock size is that recorded by an accelerometer mounted at the top of the cryostat, so that it is only a qualitative indicator of the shock to the detector assembly. The ambient shock frequency indicates the incidence at that time of shocks of that size.

trivial events. Because there have been no coincidences between the channels, we discuss the non-trivial events in each channel in turn. The distribution of events is shown in Fig. 2.

(1) Top detector: eight non-trivial events. One of these ($-1.07 \Phi_0$) occurred 2 days after initial cooldown; the cluster of six around $+1 \Phi_0$ were all in the 36 h following excitation of the calibration coil with too large a current. We believe that these seven events represent the decay of standing current in this detector loop, caused by an imperfect superconducting joint behaving as a weak link; flux changes would then be expected to occur in units of Φ_0 . We have other evidence for the presence of a weak link in this loop. On initial cooldown the current in it drifted steadily for several hours before stabilizing suddenly; a weak link with a normal state resistance of $\sim 10^{-11}$ ohm would account for the observed behaviour. These events, given that our interpretation is correct, provide an independent check of our calibration procedure: they have mean value $1.02 \pm 0.02 \Phi_0$. The rise-times of these seven events were all long, 0.5–2 s. The other event in this channel, of $0.6 \Phi_0$, had an even slower rise-time (≈ 5 s) but no evident cause. All of these events fail as candidate monopoles because of slow rise-times, incorrect magnitude, and lack of a coincident signal in the WF channel.

(2) Bottom detector: 1 non-trivial event. This event was very slow, ≈ 10 s, and had been preceded by slow drifts for several hours. It can be excluded as a monopole candidate because of the unstable background and slow rise-time, and also because of incorrect magnitude and lack of a coincident signal in the WF channel.

(3) Window frame detector: 3 non-trivial events. These three events all had bandwidth-limited rise-times, and occurred at times when the detector output was stable, without any coincident mechanical shock, pressure pulse and so on, so they at first appeared to be lacking any extraneous cause.

However, inspection of the detailed data revealed that for each of the events there was a preceding mechanical shock, as listed in Table 1. Of course, mechanical shocks do occur continually, and usually without affecting the detector. The question then is whether there is a statistically significant correlation between the shock and event. The frequency of shocks of the relevant size at the appropriate times was low on all three occasions, so that the chance of random association of the shock with the event was always $<10\%$.

In an endeavour to find a causal explanation for these events, we searched the records of putative events that had been rejected for other more obvious reasons (such as, helium level low, and concurrent pressure changes). In addition to several events with simultaneous shock, there were four events with statistically significant 'near-coincident' shocks, two of which had the shocks following the events (Table 1). In these four events, the pressure had been changing rapidly, and we believe that that caused mechanical movement giving rise to both shocks and events.

A puzzling feature of all of these 'near-coincident' shocks and events is the extended time interval, 10–140 s, between the two

(in the events we have described as being caused by coincident shocks, the shock and event have been simultaneous within the 50-ms time resolution of our data acquisition system). It seems strange too that the two largest of these putative events (numbers 34 and 35), which occurred 10 days apart, have the same size and sign.

Although we are lacking a satisfactory causal explanation of putative events numbers 34, 35 and 74, the strong circumstantial evidence of the near-coincident shocks excludes them as candidate magnetic monopoles. Both shock and event were probably caused by thermal expansion, and if we had a vibration sensor on the detector assembly itself, the explanation may have been obvious.

In conclusion, we have seen no serious candidate monopole events. Our total detector area, allowing for overlap, is 0.17 m^2 averaged over 4π , and active observation time to date is 4,850 h. This area-time product corresponds to 230 'Cabreras', the unit of exposure equivalent to that of Cabrera's original experiment. The upper limit (at 90% confidence level) from our experiment on the cosmic monopole flux is $6 \times 10^{-12} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$; taken together with the negative results reported recently (refs 9, 11, 13 and P. Chaudhari and B. Cabrera, personal communications), and disregarding Cabrera's original candidate event, the experimental upper bound on the flux now stands at $2 \times 10^{-12} \text{ cm}^{-2} \text{ sr}^{-1} \text{ s}^{-1}$, which is still three orders of magnitude greater than the Parker limit⁶.

We thank A. Cayless, J. A. Cowen, M. Koratzinos, K. Lam, D. Podger, W. S. Steer and D. Websdale for assistance. Financial support was received from the Particle Physics Committee of the Science and Engineering Research Council, and from Imperial College.

Received 26 April; accepted 10 July 1985.

1. Cabrera, B. *Phys. Rev. Lett.* **48**, 1378–1381 (1982).
2. Carrigan, R. A. Jr & Trower, W. P. (eds) *Magnetic Monopoles* (Plenum, New York, 1983).
3. Stone, J. L. (ed.) *Monopole '83'* (Plenum, New York, 1984).
4. Dirac, P. A. M. *Proc. R. Soc. A* **133**, 60–72 (1931).
5. Ahlen, S. P. in *Monopole '83'* (ed. Stone, J. L.) 629–635 (Plenum, New York, 1984).
6. Parker, E. N. in *Monopole '83'* (ed. Stone, J. L.) 125–136 (Plenum, New York, 1984).
7. Park, J. G. & Guy, C. N. in *Magnetic Monopoles* (eds Carrigan, R. A. Jr & Trower, W. P.) 208–211 (Plenum, New York, 1984).
8. Guy, C. N. in *Monopole '83'* (ed. Stone, J. L.) 125–136 (Plenum, New York, 1984).
9. Incandela, J. et al. *Phys. Rev. Lett.* **53**, 2067–2070 (1984).
10. Gardner, R., Cabrera, B., Taber, M. & Huber, M. *Proc. 17th Int. Conf. on Low Temperature Physics* 945–946 (North-Holland, Amsterdam, 1984).
11. Tesche, C. D., Chi, C. C., Tsuei, C. C. & Chaudhari, P. *Appl. Phys. Lett.* **43**, 384–386 (1983).
12. Schouten, J. C. et al. *J. Phys. E*, (submitted).
13. Cabrera, B., Taber, M., Gardner, R. & Bourg, J. *Phys. Rev. Lett.* **51**, 1933–1966 (1983).

Association of the 6-ms pulsar PSR1953 with the COS-B γ -ray source 2CG065

P. M. Chadwick, J. C. Douthwaite, A. B. Harrison, I. W. Kirkman, T. J. L. McComb, K. J. Orford & K. E. Turver

Department of Physics, University of Durham, Durham DH1 3LE, UK

A systematic search for fast radio pulsars in the error boxes of the low energy γ -ray sources detected by the COS-B collaboration led to the discovery¹ of a 6.1-ms radio pulsar PSR1953+29 in a binary system with orbital period of 120 ± 4 days inside the error box of 2CG065+0. It was not possible to associate unambiguously the two objects without either a precise position for the γ -ray source or the detection of 6.1-ms pulsed γ rays. Following the determination of a more accurate orbital ephemeris for the radio pulsar, we have examined our data for evidence for pulsed TeV γ rays. Here, we report the detection of such γ rays with the characteristic 6.1-ms periodicity of the radio pulsar. The γ rays were detected using the ground-based atmospheric Cerenkov technique at the University of Durham γ -ray facility at Dugway, Utah². This result, significant at the 5.4σ level, provides evidence for the association of the 6-ms radio pulsar PSR1953+29 with the γ -ray source 2CG065+0.

Table 1 The constants of PSR1953+29

RA:	19 h 53 min 26.673 s
Dec:	29° 0' 44.1"
Epoch:	2445428.66
<i>d</i> :	117.3 ± 0.1 days
<i>e</i> :	<0.001
<i>p</i> :	6.133166 ms
asin(<i>i</i>):	31.29 light s

Data from R. Buccheri, personal communication.

Observations of PSR1953+29 were made between 4 July and 10 October 1983 and during July 1984 using the array of four telescopes of the Dugway VHE γ -ray facility which have been described elsewhere². The telescopes have an energy threshold dependent on zenith angle of ~ 1 TeV. Three of them have an aperture of $\sim 2^\circ$ FWHM, whilst the fourth has an aperture of about 1.3° FWHM. The duration of the 20 individual observations varied between 1 and 6 h during which the pulsar was observed with all telescopes in the tracking mode. The time of arrival of a Cerenkov light flash was recorded in Coordinated Universal Time (UTC) to an absolute accuracy of 0.5 ms and with a resolution of $1 \mu\text{s}$. The system clock was based on an ovened crystal checked against off-air timing transmissions and had a stability of better than 1 part in 10^9 . This allowed reliable linking of events with a 6-ms period in phase during a single observation but not during the total exposure spanning a year. The time of each event has been converted from UTC to Barycentre Corrected Julian Ephemeris Time using the MIT Solar System ephemeris³ and the position of the radio pulsar. A further correction to allow for the orbital motion of the pulsar has been made by adjusting the times to the focus of the binary orbit according to the radio measurements summarized in Table 1.

A homogeneous dataset was compiled comprising only those eight observations of at least 4 h duration. This enables the results of the periodicity searches on individual observations to be combined with roughly equal statistical weights. The uncertainty in our knowledge of the binary orbit and the stability of the system clock causes progressive phase errors and precludes the use of datasets of 24 h or more, in particular the assembly of a single large dataset spanning the whole observation period. Initially, events were selected that triggered any two of the four telescopes (this has been shown^{4,5} to increase the sensitivity of the telescope array to γ rays as twofold telescope responses define a narrower effective aperture). The dataset of twofold telescope responses (comprising 14,286 events in the 8 observations) was obtained during the half of the orbit when the pulsar was approaching the Earth, but spanned a year, or approximately three orbits. Each individual observation of 4 h or longer was tested for periodicity over a restricted range of period (6.133162–6.133170 ms). This limited searching for periodicity was necessary to allow (with decreasing effect) for: (1) uncertainties in the orbital ephemeris; (2) the effects of statistical sampling on a true periodicity in sparse data; and (3) residual uncertainties in the precise rate of the stabilized 1 MHz system clock. The Rayleigh test^{5,6} was appropriate for searching for periodicity. In the absence of specific knowledge, or a reasonable guess, of the light-curve duty cycle this is the most general test and gives the probability that the fundamental of the trial periodicity is present, and if present its power. The test is 'uniformly most powerful' for distinguishing between the alternative hypotheses of the fundamental power being zero or non-zero. We consider that the use of such a general test for the initial scanning of a dataset for possible periodicity, although it may underestimate the total signal strength, is the most conservative.

Each of the eight observations was analysed independently using the Rayleigh test and the eight probabilities of chance occurrence of each of a fixed set of trial periods were combined⁷. This combined probability is shown as a function of the trial

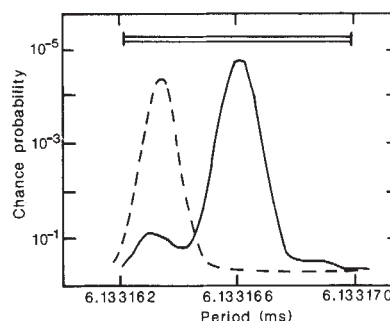


Fig. 1 Probability of chance occurrence as a function of the trial period. Horizontal bar, expected position of the radio period (see Table 1) allowing for the uncertainties in the binary orbit; solid line, two-telescope responses; broken line, independent sample from the single (sensitive) telescope response.

period in Fig. 1. A $3.5 \pm 0.9\%$ periodic excess is found within the trial period range, which has a probability of chance occurrence of 1.6×10^{-5} . The number of independent trial periods in the range: $P_1 = 6.133162$ ms to $P_2 = 6.133170$ ms is $(T/P_1 - T/P_2) = 3.4$ (T = mean duration of an observation), increasing the probability of the detection being the result of chance to 5.4×10^{-5} .

Further totally independent data from the exclusive single-telescope responses during the same eight observations is available. Such single-telescope responses are less sensitive to γ rays because of the larger aperture which allow the registration of more proton-induced showers. However, one of the four telescopes had been equipped with new mirrors and has a narrower field of view to improve its signal/noise ratio to approximately the same as two-telescope responses. This totally independent dataset, comprising 17,302 events none of which are considered in the twofold set, shows a $3.2 \pm 0.8\%$ pulsed signal at a period within the sampling range of the periodicity shown by the two-telescope responses. This signal is at a significance level of 4×10^{-5} (see Fig. 1). With a signal strength of $\sim 3.5\%$ in narrow aperture detection systems we would expect the single-telescope responses of the other three wider aperture telescopes to have a signal/noise ratio of about 1.5% and hence not show significant effects. This is seen to be the case. The overall probability of the effect being of chance origin is therefore about 4×10^{-8} which corresponds to the detection of a 5.4σ signal.

The γ -ray signal has been investigated on a night-to-night basis during the eight longest observations. The average signal (γ -rays)/noise (cosmic-ray protons in the field of view) for a two-telescope response is 3.5% and fluctuates from below the background noise level to 8.5% . No pattern is apparent in the variation, in that the signal averaged over observations which are at similar orbital phases remains at $\sim 3.5\%$, although in some cases the constituent observations are a year apart. Assuming that the γ -ray phases have a Von Mises distribution, we can test for constancy of the signal from night to night using an established test⁶. We find that the variation of signal strength from night to night is consistent with sampling variations from a constant signal strength, but cannot preclude variations from zero to about three times the average strength. We can, however, exclude the origin of the detected γ rays in strong ($\sim 30\%$) but brief (of a few minutes duration) outbursts as detected by us from the Crab pulsar⁵ and from Hercules X-1⁸. The time-averaged flux of γ rays of energy > 2 TeV, the energy threshold for the two-telescope responses, is $(3 \pm 0.8) \times 10^{-11} \text{ cm s}^{-1}$. The very high energy (VHE) γ -ray luminosity, assuming a distance to the pulsar of 3.5 kpc and a differential spectral slope of 3.0 is $(3 \pm 0.8) \times 10^{35} \text{ erg s}^{-1}$.

It is not possible to maintain phase from night to night for such a fast pulsar in an imprecisely known binary orbit, and thus it is not possible to derive an overall light curve for the dataset. The light curve for the two-telescope responses on the occasion of the apparently strongest signal (6 August 1983) has

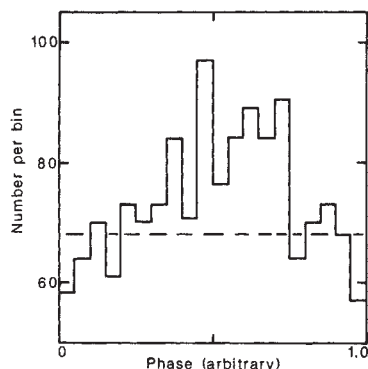


Fig. 2 Light curve of VHE γ -ray emission on 6 August 1983.

been derived for the radio period (6.133166 ms) and is shown in Fig. 2.

It is interesting to note that the three binary VHE γ -ray pulsars so far discovered by us (X-ray pulsars Her X-1⁸, 4U0115+63⁹ and radio pulsar PSR1953+29) have light curves with a very broad pulse ($\sim 30\%$ width), although they differ greatly in their pulsar periods and companion characteristics. This is in contrast to the very narrow ($\sim 1\%$ width) peaks we have seen in the VHE γ -ray light curve of an isolated radio pulsar, PSR0531+23^{4,10}. If this connection is borne out by observations of further pulsars, then either the magnetic fields and spin axes of pulsars in binary orbits are almost aligned, or the basic VHE γ -ray production mechanisms are very different for the two classes of γ -ray pulsar. A common model for the emission of isolated

pulsars involves electrons and positrons produced and accelerated in the outer magnetospheric gap very near to the radius of light cylinder, the narrow peak being caused by the small tolerance in the angle the VHE electrons make with the magnetic field lines. The binary pulsars must obviously be relieved from this constraint in some way to give 30–40% width peaks. The sporadic nature of the VHE γ -ray emission of the slow binary pulsars points to the involvement of accretion in either altering the emission mechanism itself, or in the shielding or reprocessing of VHE γ rays produced by the normal mechanism. In the latter case, the maximum γ -ray energy on production would be $\gg 1$ TeV. Note that the Cyg X-3 binary system, observed by us and other groups² to emit 1 TeV γ -rays (not as yet detected as a pulsar), also emits 1 PeV γ -rays^{12,13} modulated by the binary orbital period.

This work was funded by the SERC and carried out at Dugway Proving Ground, Utah. We thank the commander and staff of Dugway Proving Ground who supported our experimental programme and Professor R. Buccheri for the data on the radio pulsar.

Received 23 April; accepted 2 August 1985.

1. Boriakoff, V., Buccheri, R. & Fauci, F. *Nature* **304**, 417–419 (1983).
2. Dowthwaite, J. C. *et al. Astr. Astrophys.* **126**, 1–6 (1983).
3. Ash, M. E., Shapiro, I. I. & Smith, W. B. *Astron. J.* **72**, 338–350 (1967).
4. Dowthwaite, J. C. *et al. Astrophys. J.* **286**, L35–L38 (1984).
5. Gibson, A. I. *et al. Nature* **296**, 833–835 (1981).
6. Mardia, K. *Statistics of Directional Data* (Academic, London, 1972).
7. Eadie, W. T., Drijrd, D., James, F. E., Roos, M. & Sadoulet, B. *Statistical Methods in Experimental Physics* **283**, (North-Holland, Amsterdam, 1971).
8. Dowthwaite, J. C. *et al. Nature* **309**, 691–693 (1984).
9. Chadwick, P. M. *et al. Astr. Astrophys. Lett.* (in the press).
10. Chadwick, P. M. *et al. Astr. Astrophys. Lett.* (in the press).
11. Samorski, M. & Stamm, W. *Astrophys. J. Lett.* **268**, L17–L21 (1983).
12. Lloyd-Evans, J. *et al. Nature* **305**, 784–787 (1983).

Physical state of volatiles on the surface of Triton

Jonathan I. Lunine* & David J. Stevenson†

* Lunar and Planetary Laboratory, University of Arizona, Tucson, Arizona 85721, USA

† Division of Geological and Planetary Sciences, California Institute of Technology, Pasadena, California 91125, USA

The most recent analyses of infrared spectrophotometric studies of Neptune's satellite Triton concluded that both condensed methane and nitrogen are present^{1,2}. It was also concluded² that the most likely surface configuration is a liquid nitrogen (N_2) 'ocean' with dry areas of solid methane (CH_4), and perhaps some exposed fine-grained water frost (H_2O). However, this model runs into some difficulties, especially when requirements of phase equilibrium between the solid and liquid components are imposed. Because an understanding of the distribution and state of volatiles is crucial in interpreting secular changes in Triton's appearance due to seasonal effects³, and in planning observing strategies for the Voyager–Neptune/Triton encounter, we assess here several possible configurations for these volatiles on Triton. We conclude that the simplest volatile configuration which best satisfies the constraints with the least number of *ad hoc* assumptions is N_2 and CH_4 both in solid forms, perhaps partly as a microscopic mixture, but more probably as a disequilibrium assemblage, non-uniformly distributed. Thermodynamic equilibrium is then limited by seasonal transport and the finite diffusion time of CH_4 in crystalline N_2 . Although a nitrogen ocean cannot be excluded, it requires very restrictive assumptions.

In Table 1, we list five model surface configurations, and we have selected five criteria against which each model is tested. Criteria 1–4 are based on the analyses of the near-infrared spectra. Cruikshank and Apt¹ conclude that the methane absorption bands are due primarily to the solid phase, with a very small contribution from the gas (criterion 1); moreover,

Table 1 Tests of Triton models

	Criterion				
	1	2	3	4	5
(A) CH_4 -saturated N_2 ocean + solid CH_4 patches	(×)	✓	×	✓	(✓)
(B) Nearly-pure N_2 ocean + solid CH_4 patches	✓	✓	✓	✓	×
(C) N_2 ocean + small quantity of dissolved CH_4	(✓)	(✓)	✓	✓	(✓)
(D) Greenhouse N_2 atmosphere + N_2/CH_4 clouds	✓	(✓)	(×)	(×)	✓
(E) Solid N_2 + solid CH_4 and/or mixed solid N_2-CH_4	✓	✓	✓	✓	✓

✓, Model consistent with criterion; ×, model inconsistent with criterion; (), Insufficient data to rigorously determine consistency.

The criteria are:

- (1) Shape of CH_4 absorption features in spectrum.
- (2) Variability of CH_4 absorption features as a function of orbital aspect.
- (3) Presence of N_2 feature in spectrum.
- (4) Presence of additional absorption, possibly water ice or dissolved water.
- (5) Phase equilibria in N_2-CH_4 system.

Cruikshank *et al.*² require that the grain size be limited to $\sim 10 \mu m$. The absorption features are variable as a function of Triton orbital aspect¹; hence, the observable methane is non-uniformly distributed over the surface (criterion 2). A feature at $2.16 \mu m$ is interpreted to be the first overtone of a density-induced vibration rotation transition in N_2 (ref. 2). In our discussion, we accept the attribution of the $2.16 \mu m$ feature which is not seen by all observers⁴. Although Cruikshank *et al.*² fit the Triton spectrum with the transition in liquid N_2 and a 200-cm pathlength, the more general requirements (criterion 3) are a sufficiently high pathlength in condensed N_2 , or high pressure in the vapour phase plus long pathlength. The presence of residual absorption features suggests either a contaminant (most likely H_2O) in the N_2 or CH_4 , or H_2O 'bedrock' exposed on the

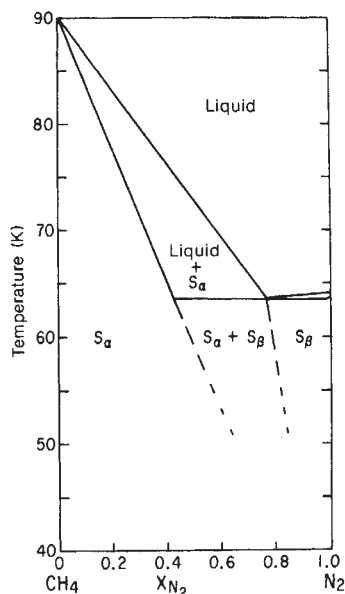


Fig. 1 Polybaric phase diagram of CH_4 and N_2 , plotted as temperature versus N_2 mole fraction, from ref. 6. S_α , cubic solid; S_β , hexagonal solid. Solid-solid boundaries are uncertain.

surface²; we adopt the requirement (criterion 4) that if H_2O is present, then its absorption is $\leq 1 \mu\text{m}$ equivalent thickness.

Criterion 5 requires that the configurations of N_2 and CH_4 coexisting on the surface are consistent with phase equilibrium studies^{5,6} or, if not, that disequilibrium is consistent with expected seasonal transport effects or kinetic considerations in the solid phase. Figure 1 shows the 'polybaric' N_2 - CH_4 phase diagram for solid-liquid equilibria^{5,6}, that is, the condensed phases are confined against their own vapour pressure. The system exhibits a nearly degenerate eutectic at $\sim 62.5 \text{ K}$ and $\sim 24\% \text{ CH}_4$ ($X = 0.24$) (ref. 6), essentially the same temperature as the pure N_2 freezing point. Materials which can lower the N_2 freezing point, such as neon, are not likely to be abundant on Triton⁷. Mixed crystal phases of N_2 and CH_4 can occur, with substitution of N_2 in the cubic (α) structure of solid CH_4 or substitution of CH_4 in the hexagonal (β) lattice of N_2 for low CH_4 concentration.

The temperature regime over which N_2 -rich liquid and N_2 - CH_4 solid coexist is within 0.5 K of 63 K on the N_2 -rich side of the eutectic^{5,6}. Thus, any N_2 liquid with $X < 0.24$, in thermodynamic equilibrium with CH_4 in a solid phase must have a temperature within 0.5 K of 63 K . Because we can think of no feedback mechanism that would cause Triton to prefer this state over others, it has a low *a priori* probability of occurring; therefore, coexistence of solid and liquid on Triton implies $X \geq 0.24$.

Models A to C involve a N_2 liquid ocean which is preferred by Cruikshank *et al.*². Model A postulates areas of dominantly- CH_4 solid in coexistence with liquid N_2 . The CH_4 solid (which could float on the N_2 - CH_4 ocean) provides the variability observed in the CH_4 absorption as a function of Triton's orbital aspect, and, if the particle size were small enough ($\sim 10 \mu\text{m}$), reproduces the shape of the methane absorption features. Arguments presented above require that such an ocean must have $T > 63 \text{ K}$. An N_2 liquid evaporating at this temperature would lose energy by evaporation at a rate 10^7 times that due to direct radiation; the corresponding factor for solid CH_4 sublimation is 10^4 . This implies that the solar insolation impinging on these volatiles is areally redistributed by sublimation. In the limiting case that the N_2 ocean is global, it can be shown³ that an isothermal atmosphere must exist which radiates uniformly over 4π . For an albedo of 0.4 , $T = 45 \text{ K}$ for unity emissivity ϵ and 63 K for $\epsilon \sim 0.3$. A global liquid ocean can then exist only if the N_2 and CH_4 possess low effective infrared emissivities. If the volatiles are restricted to polar regions, Trafton³ has shown

that the cap temperatures must be uniform and less than 63 K unless $\epsilon \leq 0.5$. Equatorial patches of volatiles are stable only during extreme solstice conditions when the diurnally averaged equatorial solar flux is negative³. Hence, the presence of liquid N_2 hinges on that material having a low emissivity.

Data on the frequency-dependent infrared emissivity of cryogenic N_2 and CH_4 liquids⁸, when weighted by the Planck function appropriate to Triton, yield an optical depth τ_{IR} of unity over a pathlength $\leq 2 \text{ cm}$. Conversely, the absorptivity of CH_4 liquid (and by inference CH_4/N_2 solid and liquid) in the visible is very low⁹, and unity optical depth to sunlight, τ_v , is obtained only over pathlengths $> 1 \text{ m}$. Thus, $\tau_v \ll \tau_{\text{IR}}$ for the Triton surface volatiles, resulting in an increasing temperature from the surface, down to the level at which the bulk of the solar flux is absorbed. Assuming an effective temperature of 50 K , a radiative temperature profile reaches the N_2 melting point, 63 K , at $\approx 5 \text{ cm}$ depth. A conductive temperature profile from the surface inwards is less steep; $T = 63 \text{ K}$ at a depth $\geq 1 \text{ m}$. The ratio of radiative to conductive time constants, $K/(4\sigma T^3 \Delta z)$ (where K is the thermal conductivity, T the effective temperature, and Δz the layer thickness), is ~ 1 and because seasonal timescales are much greater than either time constant, the shallower conductive profile is obtained. Thus, over the metre-deep pathlengths to which the spectroscopic data are sensitive², the N_2 is likely to be solid. The presence of impurities in the ice, absorbing sunlight close to the surface, makes melting over metre-deep paths even less likely.

The large amount of dissolved CH_4 in model A violates criteria 1 and 3. Cruikshank *et al.*² require a 200-cm pathlength in the N_2 liquid to reproduce the depth of the $2.16\text{-}\mu\text{m}$ absorption band; if a solid containing $\geq 24\% \text{ CH}_4$ is in coexistence with the liquid the dissolved CH_4 pathlength is then 50 cm which in a liquid of density 1 g cm^{-3} yields an equivalent pathlength-abundance of $\approx 1/2 \text{ km-amagat}$. The effect of this substantial abundance on both the CH_4 and N_2 absorption features is difficult to assess from available data¹⁰⁻¹². Using a band model at 273 K for the $2.3\text{-}\mu\text{m}$ CH_4 band in the gas¹³, we find an equivalent width of $\approx 0.5 \mu\text{m}$ at $1/2 \text{ km-amagat}$, and an optical depth > 50 in the $2.17\text{--}2.13 \mu\text{m}$ region of the band. At 60 K , the band will be much narrower, but substantial optical depth still would be expected in the wings.

Model B circumvents this difficulty by postulating a nearly pure N_2 ocean and solid CH_4 patches, that is, a system in disequilibrium. For a globally-uniform temperature, this situation is unachievable. The vapour pressure of CH_4 over the pure solid is larger than that over the undersaturated ocean. Vapour will flow from the poles to the ocean until equilibrium is restored. A strong equator-pole temperature gradient, with CH_4 ice at the pole and the N_2 ocean at the equator, would permit such disequilibrium provided the CH_4 vapour pressure were controlled by the polar temperature. The N_2 - CH_4 solution at $\sim 70 \text{ K}$ obeys the relation $X = 0.1 - 0.01 P/P_{\text{s.c.}}$ (refs 6, 14) where P and $P_{\text{s.c.}}$ are the CH_4 vapour pressure in the ocean, and pure liquid saturation vapor pressures¹⁵. For $X = 10^{-5}$ (which gives $\approx 10\text{-}\mu\text{m}$ pathlength in the ocean) $P = 10^{-7}\text{--}10^{-6}$ bars at 70 K . If this pressure is supplied by CH_4 polar caps, cap temperatures of $45\text{--}50 \text{ K}$ are required. This strong temperature gradient between N_2 and CH_4 is not achievable in either the global ocean model or polar cap model presented above, because in either case the temperature of the volatiles, regulated by sublimation and recondensation, is uniform. Transport of volatiles seasonally also makes the physical separation of N_2 and CH_4 volatiles implausible.

Model C avoids the requirement of a temperature gradient by postulating a very low global CH_4 to N_2 mole fraction of 10^{-5} , with all CH_4 dissolved in the ocean. The CH_4 absorption bands would then vary with Triton's orbital aspect, in phase with any (as yet unobserved) variation of the N_2 absorption feature. Whether the shapes of the CH_4 absorption features on Triton can be fitted by this dilute solution has not yet been demonstrated.

Finally, A-C probably satisfy criterion 4. The solubility of

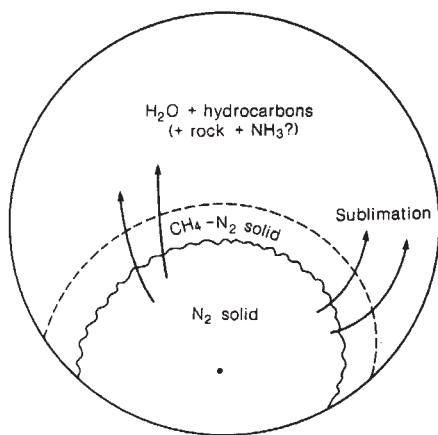


Fig. 2 Conceptual model of visible hemisphere of Triton. Nearly pure N_2 solid is present as a seasonal polar cap, and CH_4 or mixed CH_4 - N_2 solid as a lag deposit.

H_2O in cryogenic N_2 is as high as 10^{-5} mole fraction¹⁶, equivalent to a 20- μm pathlength in a 200-cm depth ocean. This is consistent with the conclusion that a small amount of H_2O absorption is present in the Triton data².

Model D postulates an atmosphere of $N_2 > 0.5$ bars, thick enough to induce a gaseous 2.16- μm absorption feature of N_2 , with a temperature gradient which permits cloud formation and hence solid CH_4 line absorptions. We assume N_2 and CH_4 to be saturated at the surface; the surface temperature is chosen to yield an N_2 vapour pressure large enough to provide a consistent pressure-induced gas opacity¹⁷ for the given effective temperature (≈ 53 K). The resulting radiative profile leads to very large surface temperatures (≈ 85 K) and massive CH_4 clouds. This model probably does not yield a sufficient pathlength in the N_2 gas to permit the 2.16- μm feature to be seen, unless clouds are avoided altogether by undersaturating both CH_4 and N_2 . It is then difficult to satisfy simultaneously criteria 1, 2 and 3.

Model E is the simplest and the most consistent with the data. N_2 and CH_4 in equilibrium below 62 K would be present as mixed crystals. To achieve a high pathlength of N_2 absorption, large ($>$ centimetre-sized) crystals of pure N_2 must be formed. Trafton³ has shown that seasonal transport of large quantities of volatiles can, in principle, occur on Triton. The higher vapour pressure of N_2 (1,000 times that of CH_4 at 50 K) ensures a distillation effect in which N_2 is transported in preference to CH_4 , resulting in a heterogeneous distribution on the surface which may involve nearly pure N_2 in some regions. Assuming the mass flux of N_2 is driven by the excess of solar flux, at the subsolar point, over that required to maintain the global (or polar cap³) temperature average, we estimate that ≈ 1 m of N_2 can be sublimed and redeposited over 10^2 yr ($\epsilon = 1$, $A = 0.4$). The unusual seasonal behaviour of Triton ensures that some N_2 may remain in place for $\approx 10^2$ yr. Disequilibrium between N_2 and CH_4 is much more likely if both volatiles are in the solid phase because of the probably slow diffusion rate of N_2 and CH_4 , ~ 1 cm over 10^2 yr based on a diffusion model¹⁸.

Grain growth calculations based on semi-empirical models¹⁹ suggest pure N_2 grains can grow up to 10 cm during this time; this has not been verified experimentally. Although the 2.16- μm absorption feature has not been measured in N_2 solid², it should be nearly identical to that in the liquid near the solidus, as is demonstrated by other Van der Waals-bonded solids¹⁰. The CH_4 ice can grow to 0.1 cm over 10^2 yr; however, most of the CH_4 could be present as a lag deposit in mixed CH_4 - N_2 crystals which could limit the CH_4 path length. Exposure of fine-grained water ice in equatorial regions is likely in any model in which the N_2 and CH_4 volatiles are solid³ as is the coating of CH_4 -derived polymers in areas devoid of volatiles. Figure 2 provides a conceptual scenario for the present, visible face of Triton.

We conclude that model E, involving CH_4 and N_2 solids, is consistent with all criteria discussed, and requires a minimum of restrictive assumptions. Observations of diurnal and seasonal variations in N_2 and CH_4 absorption features as summer solstice nears³ is of high priority in testing this conclusion. Understanding of the volatile budget on Triton, which bears strongly on cosmochemical issues⁷, as well as planning of Voyager 2 observations, requires a baseline model of the surface state of Triton's volatiles. Also of interest are the detection of argon (Ar) and carbon monoxide (CO). Argon is intermediate in volatility between N_2 and CH_4 (ref. 15) and forms solid solutions with both^{20,21}. CH_4 dissolved in Ar solid exhibits absorption features shifted from the pure ice²². While the physical chemistry of CO is very similar to that of N_2 (ref. 23) the permanent dipole moment of CO produces spectral features which have not as yet been distinguished from CH_4 bands in the Triton data². These mixed crystal compounds may have interesting and detectable coupled modes which should be investigated as important future diagnostics of the nature of Triton's surface.

This work was supported by NASA grants NAGW-185 and NAGW-450 and NSF grant AST-8206173. Contribution no. 4236 from the Division of Geological and Planetary Sciences, California Institute of Technology.

Received 26 March; accepted 8 July 1985.

1. Cruikshank, D. P. & Apt, J. *Icarus* **58**, 306-311 (1984).
2. Cruikshank, D. P., Brown, R. H. & Clark, R. N. *Icarus* **58**, 293-305 (1984).
3. Trafton, L. *Icarus* **58**, 312-324 (1984).
4. Rieke, G. H., Lebofsky, M. J. & Lebofsky, L. A. *Nature* (in the press).
5. Moran, D. W. thesis, Imperial College, London (1959).
6. Omar, M. H., Dokoupil, Z. & Schrotten, H. G. M. *Physica* **28**, 309-329 (1962); Connolly, D., Milbrodt, T. & Stoner, R. J. *J. chem. Phys.* **73**, 5388-5390 (1980).
7. Lunine, J. I. thesis, California Institute of Technology, Pasadena (1985).
8. Jones, M. C. *Far Infrared Absorption in Liquified Gases* (NBS Techn. Note 390, Washington DC, 1970).
9. Patel, C. K. N., Nelson, E. T. & Kerl, R. J. *Nature* **286**, 368-370 (1980).
10. Ewing, G. E. *J. chem. Phys.* **40**, 179-183 (1964).
11. Bulunin, M. O. & Melnik, M. G. *Opt. Spectrosc.* Suppl. **3**, 110-113 (1968).
12. Kwok, J. & Robinson, G. W. *J. chem. Phys.* **36**, 3137-3140 (1962).
13. Cruikshank, D. P. & Silvestro, P. M. *Astrophys. J.* **233**, 1016-1020 (1979).
14. Lunine, J. I. & Stevenson, D. J. in *Ices in the Solar System* (ed. Klinger, J.) (Reidel, Dordrecht, in the press).
15. Brown, G. N. Jr & Ziegler, W. T. *Adv. Cryog. Engng* **25**, 662-670 (1979).
16. Rebiai, R., Rest, A. J. & Scurlock, R. G. *Nature* **305**, 412-413 (1983).
17. Hunten, D. M. in *The Saturn System* (eds Hunten, D. M. & Morrison, D.) 127-140 (NASA CP 2068, 1978).
18. Hobbs, P. V. *Ice Physics*. 378 (Clarendon, Oxford, 1974).
19. Clark, R. N., Fanale, F. P. & Zent, A. P. *Icarus* **56**, 233-245 (1983).
20. Van 'T. Zelfde, P., Omar, M. H., Le Pair-Schrotten, H. G. M. & Dokoupil, Z. *Physica* **38**, 241-252 (1968).
21. Long, H. M. & DiPaolo, F. S. *Chem. Engng. Prog. Symp. Ser.* **59**, issue 44, 30-35 (1963).
22. Cabana, A., Savitsky, G. B. & Horning, D. F. *J. chem. Phys.* **39**, 2942-2950 (1963).
23. Brown, B. R. in *Mellors Comprehensive Treatise on Inorganic and Theoretical Chemistry* Vol. 8 (Suppl. 1) 27-149 (Wiley, New York, 1964).

Influence of transition metal complexes on atmospheric droplet acidity

T. E. Graedel*, C. J. Weschler† & M. L. Mandich*

* AT&T Bell Laboratories, Murray Hill, New Jersey 07974, USA

† Bell Communications Research, Holmdel, New Jersey 07733, USA

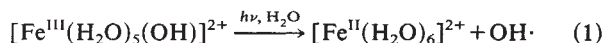
Transition metals are common constituents of cloud droplets, raindrops and other atmospheric droplets¹⁻⁵ (see Table 1). To evaluate the relative importance of pathways catalysed by transition metals in atmospheric droplet chemistry, especially reactions that increase the net acidity of atmospheric droplets, we have now formulated the first detailed model of such chemistry which incorporates transition metal complexes. Results from our computer simulations suggest that at pH 4, transition metal catalysed pathways account for 30-55% of the oxidation of S(IV) (sulphur dioxide and related sulphites) to S(VI) (sulphuric acid and related sulphates). Furthermore, the calculations suggest that transition metal reactions are the principal sources of free radicals in atmospheric droplets and that a catalytic cycle involving the Fe(II)-Fe(III) couple is an efficient producer of organic acids from oxidation of the ubiquitous atmospheric aldehydes.

Table 1 Comparison of transition metal concentrations in rain from different regions

Location	Concentration ($\mu\text{g l}^{-1}$)				Ref.
	Fe	Cu	Mn	Ni	
Minnesota/Wisconsin	5–1,400	5–150	5–60	30–240	1
Britain	130–3,100	23–82	12–23	3.8–23	2
Bermuda	1.9–13	0.1–1.6	0.1–1.0	0.04–1.4	3
North/South Atlantic	50–450	5–550	1–12	<2.5–19	4
Enewetak Atoll	0.2–2.7	0.005–0.15	0.01–0.3	—	5

Transition metals dissolved in aqueous solution exist as complexes with other solution constituents (see Table 2). For Mn(II), Cu(II) and Ni(II), given typical concentrations of possible ligands in atmospheric droplets and reported stability constants between these metals and the various ligands, the hexaquo complexes (in which the metal ion is bound to six water molecules) are the overwhelmingly dominant species. The situation for Fe(III) is more complicated; it is distributed between $[\text{Fe}(\text{H}_2\text{O})_6]^{3+}$, $[\text{Fe}(\text{H}_2\text{O})_5(\text{OH})]^{2+}$ and $[\text{Fe}(\text{H}_2\text{O})_4(\text{OH})_2]^+$, the distribution being governed by the solution pH. Furthermore, the stability constants for the Fe(III) complexes with sulphite and sulphate are sufficiently large that complexes with these ligands can exist in equilibrium with the aquo and hydroxide complexes.

While all the above transition metal complexes absorb significantly in the ultraviolet region of the spectrum, the complexes of Fe(III) have absorptions that extend into the 290–400 nm region, overlapping the solar spectrum (Fig. 1). On the absorption of a solar photon, the Fe(III) complexes undergo charge transfer from the ligand to the metal centre^{6,7}. The net result is one electron reduction of the metal centre, coupled with the oxidation of an electron donor (the coordinated ligand). The redox process is followed by the escape of the donor ligand radical into the bulk solution and solvation of the resulting Fe(II) complex:



The photochemistry of the Fe(III) hexaquo, sulphite and sulphate complexes are similar, but their absorption cross-sections between 290 and 400 nm are at least an order of magnitude less. As a result, the photoreactions of the hydroxide complexes are those of particular chemical relevance⁸. (Note that photo-

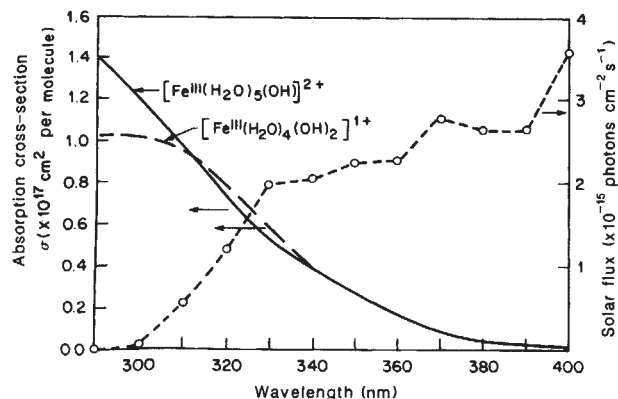
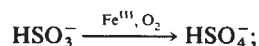


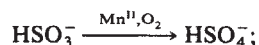
Fig. 1 Absorption spectra of $[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_5(\text{OH})]^{2+}$ and $[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_4(\text{OH})_2]^+$ in the region 290–400 nm. The absorption cross-sections for $[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_5(\text{OH})]^{2+}$ are obtained from ref. 31; those for $[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_4(\text{OH})_2]^+$ were measured in this work using standard spectrophotometric techniques. At wavelengths between 340 and 400 nm, the spectra for these two complexes overlap, and separate lines are not shown. The solar spectrum refers to a vertical height of 0.75 km (halfway in the 1.5-km fall to ground used in our model), and a north latitude of 41° at noon on 1 May. The circled points indicate calculated averages (over 10-nm intervals) of the incident solar flux for clear skies (ref. 32), reduced by the presence of the model cloud. The dotted line through these points is intended only as a guide for the reader.

oxidation of the Fe(II) hexaquo complex, the principal Fe(II) species, is negligible over the tropospheric solar spectrum.)

Our model calculations are designed to simulate chemical processes within raindrops falling from cloud base to the ground. Some portions of the reactants are within the drops prior to their fall, others are incorporated during descent. The reactions and calculations are based in part on our previous work^{9,10}, but include nearly 50 reactions involving Fe(II), Fe(III), Mn(II), Mn(III), Cu(I) and Cu(II) as well. In particular, reactions (1) and (2) are incorporated, as are the potential S(IV) oxidizing reactions involving iron and manganese¹¹.



$$\frac{-d[\text{S(IV)}]}{dt} = 0.82[\text{S(IV)}][\text{Fe}^{\text{III}}]/[\text{H}^+] \text{ Ms}^{-1} \quad (3)$$



$$\frac{-d[\text{S(IV)}]}{dt} = 25.0[\text{S(IV)}][\text{Mn}^{\text{II}}]/[\text{H}^+] \text{ Ms}^{-1} \quad (4)$$

The total concentrations of soluble transition metal ions (distributed appropriately into the various complexes) are taken to be 2×10^{-5} M for iron, 2×10^{-6} M for manganese and 2×10^{-6} M for copper, well within measured limits (see Table 1).

An essential ingredient to these calculations is the quantum yield, Φ , for the photochemical production of radicals as in reactions (1) and (2). Measurements of Φ for $\text{OH} \cdot$ radical production from Fe(III) complexes in cloud or rain water have not been performed. There have been, however, numerous laboratory studies of the photochemistry of aqueous Fe(III) species. In applying the measurements of Φ from these experiments to our model system, we have taken into account the effects of pH, the presence of ionic species and oxygen^{8,12,13}, overall Fe(III) complex concentrations, ensuing Fe(II) concentrations, and irradiation wavelength. The result is a conservative estimate of 2% for the wavelength independent Φ for $\text{OH} \cdot$ radical production. Actual values of Φ may be as much as an order of magnitude higher.

Calculations were carried out at pH 3 and pH 4, for both day and night conditions. Table 3 shows the results for S(IV) oxidation. At pH 4, reactions (3) and (4) account for 30–35% of this oxidation, the remainder being due to H_2O_2 . At pH 3, H_2O_2

Table 2 Calculated upper limit concentrations for the major metal complexes in atmospheric droplets at selected pH values

Metal complex	Concentration (M)		
	pH 3	pH 4	pH 5
$[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_6]^{3+}$	$2 \times 10^{-6*}$	—	—
$[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_5(\text{OH})]^{2+}$	$1 \times 10^{-5*}$	$3 \times 10^{-6*}$	$5 \times 10^{-8*}$
$[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_4(\text{OH})_2]^+$	$3 \times 10^{-6*}$	$1 \times 10^{-5*}$	$2 \times 10^{-6*}$
$[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_5(\text{SO}_3)]^+$	$2.2 \times 10^{-6*}$	$6 \times 10^{-6*}$	$1 \times 10^{-6*}$
$[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_4(\text{OH})(\text{SO}_3)]$	—	$6 \times 10^{-7*}$	$1 \times 10^{-7*}$
$[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_5(\text{SO}_4)]^+$	$1 \times 10^{-6*}$	—	—
$[\text{Cu}^{\text{II}}(\text{H}_2\text{O})_6]^{2+}$	2×10^{-6}	2×10^{-6}	2×10^{-6}
$[\text{Ni}^{\text{II}}(\text{H}_2\text{O})_6]^{2+}$	2×10^{-6}	2×10^{-6}	2×10^{-6}
$[\text{Mn}^{\text{II}}(\text{H}_2\text{O})_6]^{2+}$	2×10^{-6}	2×10^{-6}	2×10^{-6}

Metal complex concentrations calculated using published³⁰ values for stability constants and typical concentrations reported for ligands (for example, SO_4^{2-} , NO_3^- , Cl^-) in atmospheric droplets. A dash indicates that the species is not a significant complex at the indicated pH.

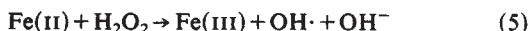
* Based on 2×10^{-5} M for total dissolved iron species in precipitation.

† Based on a Ksp limit¹⁰ of 7.9×10^{-11} M for the concentration of $[\text{Fe}^{\text{III}}(\text{H}_2\text{O})_6]^{3+}$ at pH 5; concentrations of other Fe(III) species determined by equilibrium considerations.

reaction dominates S(IV) oxidation, in part because the transition metal oxidation rates show inverse pH dependence¹¹.

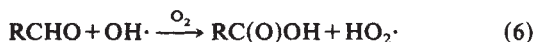
Our results also indicate previously unsuspected involvement of transition metal reactions in the organic chemistry of droplets. This is a consequence of the production of OH radicals from the transition metal reactions. The computations suggest that production of the OH· radical by reactions (1) and (2) may be the primary daytime source of solution radicals, exceeding the flux of OH· by incorporation of gas phase radicals¹⁴ even for iron ion concentration as low as 10⁻⁷ M.

At night, the photochemical reactions (1) and (2) are inoperative, and the primary source of free radicals becomes the classic Fenton reaction¹⁵



In these conditions, it turns out that the [S(IV)]/[H₂O₂] concentration ratio is an important parameter for solution oxidation. If [S(IV)] > [H₂O₂], the H₂O₂ will be largely titrated away. Conversely, if [H₂O₂] > [S(IV)], as might occur in remote areas or in populated areas with low sulphur emission fluxes, the H₂O₂ is not completely depleted by S(IV) and some is available to form OH· radicals by reaction (5).

What is the fate of the OH· radical in atmospheric droplets? Our computations indicate that it primarily reacts with aldehydes to produce organic acids¹⁶ by

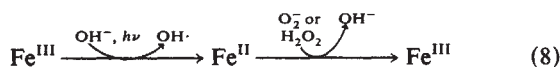


This result is of interest because of the relatively large concentrations (10⁻⁵ M or higher) of aliphatic aldehydes found in rain¹⁷⁻¹⁹, mist¹⁹, fog^{19,20} and clouds^{19,21}. Starting with initial aldehyde concentrations of 1.3 × 10⁻⁵ M, our results show organic acid concentrations of the order of 1 × 10⁻⁶ M during a 6-min fall from cloud to ground, equivalent to an oxidation rate of some 2% per min. (Much of the remaining RCHO is complexed with HSO₃⁻ (refs 21, 22) and is unavailable for oxidation; some RC(O)OH enters the droplets directly from the gas phase). Since cloud and fog droplet lifetimes are much longer than the several-minute lifetime of a raindrop one would expect extensive conversion of aldehydes to organic acids in long-lived droplets containing both H₂O₂ and transition metal ions. Indeed, concentrations of organic acids of 10⁻⁵ M or higher have been observed in precipitation on those few occasions when measurements have been made^{18,23-26} and some 16-35% of the free acidity of precipitation in the eastern United States is due to organic acids²⁷. Gas phase sources of organic acids are known, but seem insufficient to explain the organic acid concentrations found in droplets.

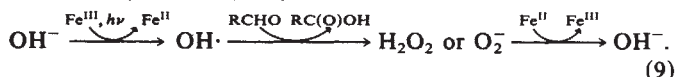
The maximum effectiveness of the transition metal processes is realized when the Fe(III) reduction (either reaction (1) or (2)) is balanced by an oxidation to produce a catalytic cycle. In atmospheric droplets this oxidation can be efficiently accomplished by O₂⁻, the conjugate base of the HO₂ produced by reaction (6):



for which a large rate constant has recently been determined²⁸. The resulting cycle has a catalytic sequence for iron of



and a complementary sequence for the O—H species of



An aspect worth noting is that O₂⁻ is in acid equilibrium with HO₂ in aqueous solution



with an equilibrium constant of pK_{HO₂} = 4.69 (ref. 29). The catalytic cycle thus becomes more efficient at higher pH values.

Therefore, our model suggests that oxidation of S(IV) in atmospheric droplets to produce S(VI) inorganic acids occurs primarily by reaction with H₂O₂ but partly by processes involv-

Table 3 Calculated percentage of S(IV) oxidation via various reactant pathways

Reactant	pH 3		pH 4	
	Day	Night	Day	Night
H ₂ O ₂	>99	>99	71	63
Fe ³⁺	<1	<1	7	8
Mn ²⁺	<1	<1	22	29

Calculations for chemical processes within a 1.0-mm diameter raindrop falling 1 km from cloud base to ground. Calculations were carried out at pH 3 and pH 4, for both day and night conditions.

ing complexed transition metal ions. A catalytic cycle involving the photoprocesses of iron(III) complexes may be important in producing organic acids from aldehydes. At night, particularly at low S(IV) concentrations, H₂O₂ fills the role of the solar photons in continuing the catalytic cycle. The rates of the reactions in this cycle depend on the concentrations of iron, H₂O₂ and RCHO, all of which are influenced by anthropogenic activities.

Received 15 March; accepted 19 July 1985.

- Pratt, G. C., Coscio, M. R. & Krupa, S. V. *Atmos. Environ.* **18**, 173-182 (1984).
- Peirson, D. H., Cawse, P. A. & Cambay, R. S. *Nature* **251**, 675-679 (1974).
- Jickells, T. D., Knap, A. H. & Church, T. M. *J. geophys. Res.* **89**, 1423-1428 (1984).
- Berresheim, H. & Jaeschke, W. *Atomispektrom. Spurenanal. Vortr. Kolloq.* (ed. Welz, B.) 305-313 (Verlag Chem., Weinheim, 1982).
- Arimoto, R., Duce, R. A., Ray, B. J. & Unni, C. K. *J. geophys. Res.* **90**, 2391-2408 (1985).
- Calvert, J. G. & Pitts, J. N. Jr *Photochemistry* (Wiley, New York, 1966).
- Endicott, J. F. in *Concepts of Inorganic Photochemistry* (eds Adamson, A. W. & Fleischauer, P. D.) 81 (Wiley, New York, 1975).
- David, F. & David, P. G. *J. phys. Chem.* **80**, 579-583 (1976).
- Graedel, T. E. & Goldberg, K. I. *J. geophys. Res.* **88**, 10865-10882 (1983).
- Graedel, T. E. & Weschler, C. J. *Rev. Geophys. Space Phys.* **19**, 505-539 (1981).
- Martin, L. R. in *SO₂, NO and NO₂ Oxidation Mechanisms: Atmospheric Considerations* (ed. Calvert, J. G.) 63 (Ann Arbor Science, Ann Arbor, Michigan, 1984).
- Adamson, A. W. *et al. Chem. Rev.* **5**, 541-585 (1968).
- Dainton, F. S. & Tordoff, M., *Trans. Faraday Soc.* **53**, 666-678 (1957).
- Chameides, W. L. & Davis, D. D. *J. geophys. Res.* **87**, 4863-4877 (1982).
- Walling, C. *Acc. chem. Res.* **8**, 125-131 (1975).
- Chameides, W. L. & Davis, D. D. *Nature* **304**, 427-429 (1983).
- Klippel, W. & Warneck, P. *Geophys. Res. Lett.* **5**, 177-179 (1978).
- Likens, G. E., Edgerton, E. S. & Galloway, J. N. *Tellus* **35B**, 16-24 (1983).
- Grosjean, D. & Wright, B. *Atmos. Environ.* **17**, 2093-2096 (1983).
- Waldman, J. M. *et al. Science* **218**, 677-680 (1982).
- Richards, L. W. *et al. Atmos. Environ.* **17**, 911-914 (1983).
- Munger, J. W., Jacob, D. J. & Hoffman, M. R. *J. atmos. Chem.* **1**, 335-350 (1984).
- Barcelona, M. J., Liljestrand, H. M. & Morgan, J. J. *Analyt. Chem.* **52**, 321-325 (1980).
- Galloway, J. N., Likens, G. E., Keene, W. C. & Miller, J. M. *J. geophys. Res.* **87**, 8771-8786 (1982).
- Keene, W. C., Galloway, J. N. & Holden, J. D. Jr *J. geophys. Res.* **88**, 5122-5130 (1983).
- Norton, R. B. *EOS - Trans. AGU* **65**, 834 (1984).
- Galloway, J. N., Likens, G. E. & Hawley, M. E. *Science* **226**, 829-831 (1984).
- Matthews, R. *Aust. J. Chem.* **36**, 1305-1317 (1983).
- Bielski, B. H. J. *Photochem. Photobiol.* **28**, 645-649 (1978).
- Smith, R. M. & Martell, A. E. *Critical Stability Constants* Vol. 4 (Plenum, New York, 1976).
- Rabinowitch, E. & Stockmayer, W. H. *J. Am. chem. Soc.* **64**, 335-347 (1942).
- Goldberg, B. & Klein, W. H. *Solar Energy* **24**, 351-357 (1980).

NMR spectrum of the potassium anion K⁻

Peter P. Edwards, Ahmed S. Ellaboudy & Dolores M. Holton

University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, UK

The conventional chemical definition of an alkali metal in solution is based on the tacit assumption of cation formation through the spontaneous ionization of the ns¹ valence electron¹. However, much experimental evidence, both direct and indirect, has accumulated over the past decade for the existence of the anionic form of the alkali elements in certain non-aqueous solutions². In 1974, Ceraso and Dye³ reported the fingerprint nuclear magnetic resonance (NMR) spectrum of the sodium anion, Na⁻, which constituted the first direct proof of the existence of this ion in solution. Since then, high-resolution multi-element NMR spectroscopy has been used to identify the ubiquitous Na⁻ ions⁴⁻⁷, as well as the rubidide

(Rb⁺) (refs 4, 7) and caeside (Cs⁺) ions⁴ in a variety of metal solution systems both liquid and solid^{8,9}. Until now, the status of the potasside ion, K⁻, has been unclear. We report here the first observation of the ³⁹K-NMR spectrum of K⁻ from solutions of potassium metal and potassium/caesium mixed metals in the liquid crown ethers, 1,4,7,10-tetraoxacyclododecane (12-crown-4, 12C4) and 1,4,7,10,13-pentaoxacyclopentadecane (15-crown-5, 15C5). This constitutes the first direct proof of the existence of K⁻; the NMR data also give detailed information on the microscopic electronic structure of the potasside anion in solution.

The alkali metals Na, K, Rb and Cs dissolve in 12C4 and 15C5 to form dark blue solutions containing metal cations, metal anions, solvated electrons and electron-cation aggregates^{10,11}. We find that mixed alkali metal samples (such as NaK and KCs) in both solvents form extremely stable solutions in which the dissolution of, for example, the mixed K/Cs metals proceeds through the disproportionation reaction



where K⁻ is the potassium anion and Cs_s⁺ is the complexed or solvated caesium cation. In these solvents, the larger alkali cations form very stable sandwich complexes of the type M⁺C₂ where C represents the crown molecule¹².

Three major problems have hindered the search for the potasside ion. The first is the intrinsic low sensitivity (receptivity) of the ³⁹K nucleus. This problem was minimized in the present work by working at the high magnetic fields (to reduce second-order quadrupolar effects) of our Bruker MSL-400 NMR spectrometer (³⁹K frequency, 18.66 MHz; 400.130 MHz proton frequency). In addition, we find that the concentration of K⁻ is enhanced by several orders of magnitude in the mixed KCs metal solutions in the liquid crowns. This had been suggested by our recent optical studies of M⁻ in these solvents (reported in detail elsewhere¹¹) and is consistent with thermodynamic considerations for the disproportionation reaction (1)¹³ for single as opposed to mixed-metal samples.

Second, solutions of the single elemental metals in 12C4 and 15C5 have high freezing points: ~283 K and 273 K respectively. Thus, relatively high observation temperatures are required to observe the liquid-state NMR spectra of K⁻. The addition of tetrahydrofuran (THF) as a co-solvent considerably reduces the freezing points of the metal solutions to ~173 K. An NMR study of liquid solutions at relatively low temperatures was, therefore, feasible. It is well established¹⁴ for metal solutions in general that solution stabilities are considerably enhanced at lower temperatures. In addition, our recent nuclear spin relaxation studies¹⁵ of Na⁻ in these crown systems suggest that NMR line broadening (through exchange processes) is not excessive in the low-temperature regime.

Third, both the electron and cation exchange processes which effectively limit the lifetime of K⁻ in solution (and consequently broaden the NMR line beyond detection) must be very inefficient if the anion resonance is to be observed. In the liquid crown systems, the exchange and dissociation equilibria of the type shown below appear to be comparatively inefficient



and



where e_s⁻ is the solvated electron.

The (exchange) line broadening of the NMR is then reduced to a minimum. Similarly, the low donaticity of the co-solvent, THF, does not appear to adversely affect the lifetime of K⁻ in these solutions.

Figure 1 shows the ³⁹K-NMR spectrum of K⁻ at 250 K from a solution of KCs/12C4 in THF. A single ³⁹K resonance is observed which is shifted to -98.2 p.p.m. (parts per 10⁶) relative to the reference signal from the aquated potassium cation, K⁺(aq) (a solution of KCl in D₂O). The large and negative chemical shift for the NMR line corresponds to a considerable increase in the nuclear shielding of the ³⁹K species compared

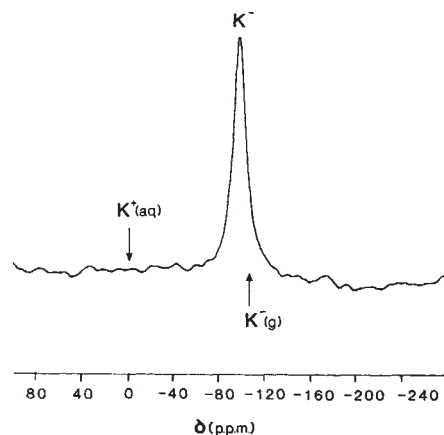


Fig. 1 The ³⁹K-NMR spectrum of a solution of KCs/12C4 in THF at 250 K. The reference (0 p.p.m.) is a solution of KCl in D₂O. A negative value of δ corresponds to an increase in the nuclear shielding of the ³⁹K nucleus. Also indicated is the calculated¹⁶ chemical shift for a gas-phase potassium anion, K⁻(g).

with ³⁹K⁺(aq) (ref. 16). This experimental NMR shift must be compared with the theoretical estimate¹⁶ of -107 p.p.m. for a gas-phase potassium anion, K⁻(g). We can, therefore, unambiguously assign the observed resonance to the potassium anion in solution. Thus, the considerable electronic shielding of the ³⁹K nucleus in K⁻ arises by virtue of the two, spin-paired 4s valence electrons.

The observed chemical shifts and NMR linewidths for Na⁻, K⁻ and Rb⁻ in 12C4 are given in Table 1; also included are the calculated chemical shifts¹⁶ for the gas-phase anions, Na⁻(g), K⁻(g) and Rb⁻(g). The data for Na⁻ in 12C4, coupled with the remarkable insensitivity of the NMR parameters to the nature of the host solvent in all systems examined^{2,4-7}, show that the sodide ion interacts only very weakly with its surroundings. To be more specific, the sodium (filled) 2p orbitals are particularly well screened from interactions with surrounding solvent or solute molecules by the presence of the filled 3s valence orbital which itself is unaffected by its environment. In addition, the extremely narrow NMR lines observed from these paramagnetic solutions provide strong evidence that Na⁻ exists as an essentially 'gas-like' moiety in solution. In contrast, the NMR of Rb⁻ in 12C4 is very significantly deshielded relative to the gaseous anion. This suggests a considerable solvent involvement in the ground-state wavefunction of Rb⁻. The deshielding of the large Rb⁻ ion is presumably caused by the introduction of orbital angular momentum through the interaction of the solvent with the rubidium 5p orbital². The broad NMR linewidth for Rb⁻ (some 1,000 Hz) is also consistent with a description of an anionic species which is significantly modified on passing from the gaseous to the condensed phase.

The NMR data for K⁻ suggests that the species is considerably more gas-like than the rubidide ion in this solvent. However, the K⁻ ion does suffer certain small perturbations on its gas-phase electronic structure. This can be gauged from the observed chemical shift of K⁻ (Table 1) which indicates a small degree of deshielding of the ³⁹K nucleus compared with K⁻(g). This is also consistent with the NMR linewidth of K⁻ which is larger

Table 1 NMR data for Na⁻, K⁻ and Rb⁻ in 12C4 solution

Ion	$\delta(\text{M}^-)_{\text{observed}}^{* \dagger}$	$\delta(\text{M}^-)_{\text{(g)}}^{* \ddagger}$	$\Delta\nu_{1/2}$ (Hz)
Na ⁻	-61.8	-63.1	1.5-10
K ⁻	-98.2	-107.0	180
Rb ⁻	-191.0	-213.6	1,000

* p.p.m. from M⁺(aq).

† Values uncorrected for bulk susceptibility corrections.

‡ Based on calculations of the nuclear shielding differences between gaseous metal anions and neutral metal atoms; see refs 4, 16.

than that of Na^+ in this solvent, but still considerably smaller than the Rb^+ value (some 1,000 Hz).

We conclude that, in the present system, the potasside ion exists as a stable, long-lived entity which interacts only weakly with its surroundings.

Finally, we note that Mendeléeff¹⁷ some 80 years ago rightly pointed out the important quantitative resemblance between the halogens and the alkali metals. With this observation of K^+ in solution, the ability of the alkali atoms to form negative ions in solution—in a manner formally akin to the halogens—is now firmly established. The only exception now is lithium, and the search for the lithium anion continues.

We thank the Science and Engineering Research Council for financial support and Drs A. E. Carpenter and J. Klinowski for technical assistance. We also thank Professor W. McFarlane for support and encouragement.

Note added in proof: Professor J. L. Dye of Michigan State University, U.S.A., has informed us that the NMR of K^+ has been independently observed from solutions of K^+ (15-crown-5)₂ K^+ in dimethyl ether¹⁸.

Received 24 June; accepted 17 July 1985.

1. Parish, R. V. *The Metallic Elements* (Longman, London, 1977).
2. Dye, J. L. *Prog. inorg. Chem.* **32**, 327–441 (1984).
3. Ceraso, J. M. & Dye, J. L. *J. chem. Phys.* **61**, 1585–1587 (1974).
4. Dye, J. L., Andrews, C. W. & Ceraso, J. M. *J. phys. Chem.* **79**, 3076–3079 (1975).
5. Edwards, P. P., Guy, S. C., Holton, D. M. & McFarlane, W. *JCS Chem. Commun.* 1185–1186 (1981).
6. Edwards, P. P. *et al. J. phys. Chem.* **87**, 4362–4363 (1983).
7. Holton, D. M. *et al. JCS Chem. Commun.* 740–741 (1984).
8. Ellaboudy, A. S., Tinkham, M. L., Van Eck, B., Dye, J. L. & Smith, P. B. *J. phys. Chem.* **88**, 3852–3855 (1984).
9. Dye, J. L. & Ellaboudy, A. *Chem. Br.* **3**, 210–215 (1984).
10. Dewald, R. R., Jones, S. R. & Schwartz, B. S. *JCS Chem. Commun.* 272–273 (1980).
11. Edmonds, R. N., Holton, D. M. & Edwards, P. P. *JCS Dalton Trans.* (in the press).
12. Poonce, N. S. & Bajaj, A. V. *Chem. Rev.* **79**, 389–445 (1979).
13. Edwards, P. P., Ellaboudy, A. S. & Holton, D. M. *Angew. Chem. Int. Ed. Engl.* (submitted).
14. Dye, C. L., Andrews, C. W. & Mathews, S. E. *J. phys. Chem.* **79**, 3065–3070 (1975).
15. Holton, D. M., Ellaboudy, A. S., Pyper, N. C. & Edwards, P. P. *J. chem. Phys.* (in the press).
16. Pyper, N. C. & Edwards, P. P. *J. chem. Soc.* (in the press).
17. Mendeléeff, D. I. *The Principles of Chemistry* 3rd edn (Longman, London, 1905).
18. Tinkham, M. L. & Dye, J. L. *J. Am. chem. Soc.* (submitted).

Hydrothermal uranium uptake at ridge crests

Annie Michard & Francis Albarede

Centre de Recherches Pétrographiques et Géochimiques et Ecole Nationale Supérieure de Géologie, F-54501 Vandœuvre Cedex, France

Although low-temperature alteration of the ocean floor traps large quantities of uranium from sea water^{1,2}, the mass balance between the crust and sea water along the mid-ocean ridges and, consequently, the uranium content of subducted lithosphere, are uncertain. Intense scavenging from sea water by particulates as a result of the dispersal of hydrothermal solutions in the water column³ has prevented the assignment of deeper levels at ridge crests as a uranium source or as a sink⁴. Here we present uranium analyses for hydrothermal solutions from the East Pacific Rise (EPR) which support earlier data^{5–6} showing that this element behaves coherently with magnesium in its removal from sea water during hydrothermal circulation. Our results indicate that oceanic crust is subducted with the U/Pb ratio increased relative to that of the ambient mantle, but this is thought to be due more to hydrothermal lead depletion than to uranium enrichment.

Since 1979, solutions from the hydrothermal vent fields on the EPR have been sampled several times at both the 13 and 21° N sites by French and American submersibles^{7–10}. At temperatures >300 °C and occasionally >350 °C, low pH, sulphide-rich and magnesium-depleted solutions are emitted in the rift valley at a depth of ~2,600 m. Evidence suggests that the hydrothermal systems are fed by sea water from the ridge flanks. The general chemistry of these hydrothermal fluids has been

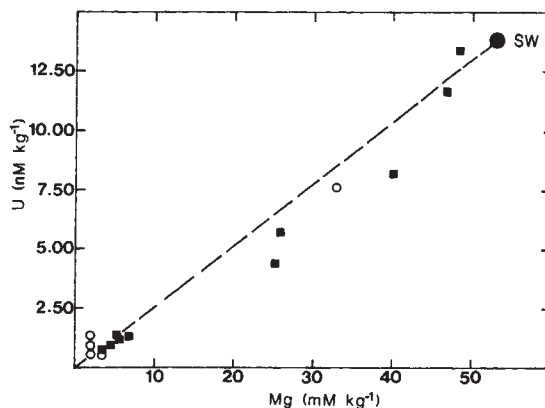


Fig. 1 Co-variation of magnesium and uranium in hydrothermal vent solutions from the East Pacific Rise. Data are taken from the present work and ref. 9. ■, 13° N; ○, 21° N.

described in refs 7–11 while preliminary data on their uranium content are given in refs 5, 6.

The present measurements on both the 13 and 21° N vent fields were carried out by isotope dilution using a combined chemistry for uranium and rare earth elements: a few millilitres of solution were spiked, purified on anion/cation exchange resins, then run on a solid-source mass spectrometer. The measurements are believed to be accurate within 2% whilst, for the least concentrated samples, blank correction amounts to a maximum of 6%.

These results are reported in Table 1, and uranium is plotted against magnesium in Fig. 1. In spite of a significant scatter, the strong positive correlation between these two elements and the quite low abundance of uranium in the samples most enriched in hydrothermal component (with the lowest magnesium concentration) suggest that both uranium and magnesium are removed from the inflowing sea water on reaction with the hydrothermal system. A few points showing intermediate magnesium content fall below the sea-water/hydrothermal component mixing line. This probably indicates incipient uranium removal by unidentified secondary minerals on mixing with sea water, as suggested for Galapagos warm vents by Edmond *et al.*³. However, such a secondary removal cannot solely account for the depletion of uranium in pure vent waters from both sites and its uptake from sea water by the hydrothermal ridge system may be considered quantitative.

Using the Wollery and Sleep¹² global estimate for the amount of water circulating within the ridge hydrothermal systems, the removal of uranium from sea water at ridge crest amounts to $5 \times 10^8 \text{ g yr}^{-1}$, a value bracketed by the previous estimates of Bloch² who correctly predicted that the deeper levels of the oceanic crust are a uranium sink. The residence time of this element in sea water relative to the hydrothermal process is identical to that of magnesium, which Edmond *et al.*¹³ estimated to be 9.6 Myr. Owing to the direction of uranium transfer, these values are insensitive to the relative importance in global hydrothermal activity of the fluids which leak out from the ridge flank at warm temperatures¹⁴.

The global uranium balance in sea water cannot be estimated with precision, due to an insufficient knowledge of the world riverine flux. Sackett *et al.*¹⁵ suggest a value of about $2 \times 10^{10} \text{ g yr}^{-1}$ while, taking into account the effects of uranium in phosphate fertilizers, Moore¹⁶ worked out a value of $1.3 \times 10^9 \text{ g yr}^{-1}$ for the pre-agricultural riverine input to the oceans. Hence, depending on the assumptions made about the riverine flux and water circulation within the oceanic crust, the ridge hydrothermal systems may balance as little as a few per cent or as much as 40% of the uranium transported to the oceans by the world river system. The strength of the sink represented by the low-temperature alteration of the uppermost oceanic crust ($0.2\text{--}1 \times 10^{10} \text{ g yr}^{-1}$ according to refs 2, 4) is also highly uncertain; this low-temperature removal of uranium in the upper

Table 1 Analytical results on submarine hydrothermal solutions from the East Pacific Rise

Site	13° N				21° N					
	Chandelier	Chainette	Cathedrale	SW 2	HG	NGS	OBS			
Sample	20 Ti4	14 Ti2	15 Ti2	20 Ti1	20 Ti3	1149-2	1157-2	1160-2	1155-10	1158-2
Mg (mM kg ⁻¹)	4.52	5.57	48.6	5.35	40	1.92	32.85	1.96	3.40	1.95
U (nM kg ⁻¹)	0.975	1.22	11.95	1.33	8.23	0.588	7.64	0.924	0.588	1.30
Pb (nM kg ⁻¹)	485	405	60	338	—	—	—	—	—	—

Mg was analysed by atomic absorption (ref. 5, 11 and courtesy of G. Michard), U and Pb were analysed by isotope dilution. Different samples may correspond to different sampling devices (Ti or graphite syringes)^{9,10}, different storage conditions (unacidified in glass vials or acidified in polyethylene flasks) and spike has been added either on board or in the laboratory. Neither contamination during sampling and storage nor adsorption on the container walls significantly affects the results.

basaltic layers may simply take place in the descending limbs of convective hydrothermal cells, which act as the recharge of the smoker systems, thereby implying that the same uranium sink is counted twice in the budget of this element.

The hydrothermal uptake of uranium from sea water also has implications for the geochemistry of the subducted oceanic lithosphere which has been proposed as the ultimate source of some ocean island basalts^{17,18}. Taking an average uranium concentration in unaltered crust as 40 p.p.m. (parts per 10⁶)⁴, 3.5 km² of oceanic crust are created annually, then 2.4 × 10⁹ g yr⁻¹ of uranium is transferred from the mantle to the new ocean crust. To this must be added the 5 × 10⁸ g yr⁻¹ of uranium derived from sea water; that is, as shown in our preliminary work⁵, the content of this element is increased by ~20% during hydrothermal alteration of the oceanic crust. The effect of this will be to increase the U/Pb ratio in altered oceanic crust.

The U/Pb ratio may also be increased by the removal of lead, again by hydrothermal reactions as shown by the presence of galena in chimneys¹⁹ and the relatively high abundance (100 p.p.b., parts per 10⁹) of this element in the EPR 21° N solutions⁶. Three solutions from the 13° N site were analysed (Table 1) with rather similar results. Redeposition of hydrothermally mobilized lead within the conduits wherever hydrothermal activity does not result in high-temperature solution venting, as on the ridge flanks, may render the extrapolation of vent data grossly incorrect. Sulphide veins in Deep Sea Drilling Project (DSDP) bore holes²⁰ attest to such redeposition. Taking the lead concentration of the fresh ocean crust to be ~100 p.p.b. (mostly concentrated in the uppermost basaltic layer), the crust produced each year would contain ~6 × 10⁹ g of lead. This quantity is significantly less than the 1.5 × 10¹⁰ g yr⁻¹ flux obtained by multiplying the presumed flux of water through the world ridge system by the lead concentration in the EPR hydrothermal solutions. This calculation is simply meant to suggest that even for an extremely poor efficiency of the hydrothermal extraction, a significant fraction of the lead contained in the newly formed crust may be dispersed into the ocean.

This leads to the question of whether uranium addition or lead extraction is the dominant factor in determining the lead systematics of basalts extracted from a source rich in recycled oceanic crust with a high $\mu(^{238}\text{U}/^{204}\text{Pb})$ ratio. In a given time interval T , if K represents the $^{232}\text{Th}/^{238}\text{U}$ ratio, a difference of:

$$\begin{aligned} & (^{208}\text{Pb}/^{204}\text{Pb})_2 - (^{208}\text{Pb}/^{204}\text{Pb})_1 \\ &= (K_2\mu_2 - K_1\mu_1) (\exp(\lambda_{232}T) - 1) \end{aligned}$$

is created between the $^{208}\text{Pb}/^{204}\text{Pb}$ ratios of the normal mantle (subscript 1) and the subducted oceanic crust (subscript 2) or basalts derived thereof. The slope of a mixing line between the two endmembers in the $^{208}\text{Pb}/^{204}\text{Pb}$, $^{206}\text{Pb}/^{204}\text{Pb}$ diagram will suggest an apparent K value which obeys:

$$(K - K_1)/(K_2 - K_1) = \mu_2/(\mu_2 - \mu_1)$$

Comparing the least and most radiogenic leads from the oceanic islands and ridges²¹ and using their $^{207}\text{Pb}/^{204}\text{Pb}$, $^{206}\text{Pb}/^{204}\text{Pb}$ 'isochron' age as the characteristic time for developing the isotopic heterogeneities results in a μ difference of almost a factor of 2 (ref. 18), much in excess of the 20% increase one

could expect from a simple hydrothermal enrichment. Hence, as suggested by Vidal and Dosso²² from other lines of evidence, the radiogenic character of the oceanic island lead results predominantly from a lead depletion in the source region of these basalts. It is not clear whether this depletion reflects a long-standing leakage of mantle lead towards the core²² or its dispersal with the hydrothermal effluents and incorporation into the sedimentary mass. However, the metalliferous sediments from a wide area straddling the EPR sequester a significant proportion of a mid-ocean ridge basalt (MORB)-like lead component²³. If these metalliferous sediments are not quantitatively returned to the mantle on subduction, or if their base metals are extracted by the volcanism associated with converging boundaries, a net flux of lead from the mantle to the crust will result.

Let us consider an extreme but simple mantle model with a two-stage evolution akin to the models considered by Chase¹⁸: the first stage involves a single μ , single K closed system evolution; at the onset of the second stage, both U/Pb and Th/U ratios are fractionated and at least two components (labelled 1 and 2) are created, the second component resulting from the first by hydrothermal alteration. The apparent source K calculated for the second stage is not independent of the Pb behaviour. When computed from the above relationship and for $\mu_2 > \mu_1$, we can expect that $K < K_1 < K_2$. In fact, an apparent value of the second stage K as low as 2.4 and a $\mu_2/(\mu_2 - \mu_1)$ value as high as 2, as Chase calculated for Hawaii¹⁸, could be obtained by assuming that K_1 is identical to the first stage value (4.1) and that the second stage K_2 is 3.25 owing to addition of 25% excess uranium during the hydrothermal alteration of the oceanic crust. It must be stressed that identical K_1 and first stage K represent no more than a convenient assumption. Obviously, this process cannot account for the apparent increase in K through time of islands such as Kerguelen¹⁸.

Financial support from the CNEXO and ATP CNRS *Géochimie-Métallologie* is gratefully acknowledged. We thank the crew and colleagues, especially the chief-scientists R. Hekinian and H. Bougault, for the success both on board R/V *Suroit* and *Nadir*, and in the darkness with *Cyana*. H. Craig shared invaluable samples.

Received 29 April; accepted 19 July 1985.

1. Aumento, F. *Earth planet Sci. Lett.* **11**, 90-94 (1971).
2. Bloch, S. *Geochim. cosmochim. Acta* **44**, 373-377 (1980).
3. Edmond, J. M. *et al. Earth planet Sci. Lett.* **46**, 19-30 (1979).
4. Hart, S. R. & Staudigel, H. *Earth planet. Sci. Lett.* **58**, 202-212 (1982).
5. Michard, G. *et al. Earth planet. Sci. Lett.* **67**, 297-307 (1984).
6. Chen, J. H., Wasserburg, G. J., Von Damm, K. L. & Edmond, J. M. *EOS* **64**, 724 (1983).
7. Edmond, J. M., von Damm, K. L., McDuff, R. E. & Measures, C. I. *Nature* **297**, 187-191 (1982).
8. von Damm, K. L., Grant, B. & Edmond, J. M. in *Hydrothermal Processes at Seafloor Spreading Centers*, NATO Conf. Ser. IV (eds Rona, P. E., Bostrom, K., Laubier, L. & Smith, K. L. Jr) 369-389 (Plenum, New York, 1983).
9. Michard, A., Albarede, F., Michard, G., Minster, J. F. & Charlou, J. L. *Nature* **303**, 795-797 (1983).
10. Grimaud, D., Michard, A. & Michard, G. *C. r. hebdom. Séanc. Acad. Sci., Paris* **299**, 865-870 (1984).
11. von Damm, K. L. Thesis, Massachusetts Inst. Technol. (1983).
12. Wollery, T. J. & Sleep, N. H. *J. Geol.* **84**, 249-276 (1976).
13. Edmond, J. M. *et al. Earth planet. Sci. Lett.* **46**, 1-18 (1979).
14. Sleep, N. H. & Wollery, T. J. *J. geophys. Res.* **83**, 5913-5922 (1978).
15. Sackett, W. M., Mo, T., Spalding, R. F. & Exner, M. E. *Symp. Interaction of Radioactive Contaminants with the Constituents of the Marine Environment*, Seattle, 757-769 (IAEA, Vienna, 1973).

16. Moore, W. S., *Earth planet. Sci. Lett.* **2**, 231-234 (1967).
17. Hofmann, A. W. & White, W. M. *Yb. Carnegie Instn Wash.* **79**, 477-483 (1980).
18. Chase, C. G. *Earth planet. Sci. Lett.* **52**, 277-284 (1981).
19. Styr, M. M. *et al. Earth planet. Sci. Lett.* **53**, 382-390 (1981).
20. Anderson, R. N. *et al. Nature* **300**, 589-594 (1982).
21. Sun, S.-S. *Phil. Trans. R. Soc. A297*, 409-445 (1980).
22. Vidal, P. & Dosso, L. *Geophys. Res. Lett.* **5**, 169-171 (1978).
23. Dash, J. E. *Geol. Soc. Am. Mem.* **154**, 199-208 (1981).

Positron emission tomography after MPTP: observations relating to the cause of Parkinson's disease

D. B. Calne*, J. William Langston†, W. R. Wayne Martin*, A. Jon Stoessel*, Thomas J. Ruth‡, Michael J. Adam‡, Brian D. Pate‡ & Michael Schulzer§

* Division of Neurology, § Department of Medicine and Statistics, UBC Health Science Centre Hospital, Vancouver, British Columbia, Canada V6T 1W5

† Department of Neurology, Stanford University School of Medicine, San Jose, California 95128, USA

‡ UBC/TRIUMF PET Program, University of British Columbia, Vancouver, British Columbia, Canada V6T 1W5

The dopa analogue 6-fluorodopa (6-FD) used with positron emission tomography (PET) allows *in vivo* visualization of dopamine and its metabolites in nigrostriatal nerve endings¹. We have now found abnormal 6-FD scans in four subjects exposed to 1-methyl-4-phenyl-1,2,3,6-tetrahydropyridine (MPTP). None had parkinsonism. The results suggest subclinical damage to the nigrostriatal pathway. This is the first direct evidence that dopaminergic impairment can exist without clinical deficits. Here we discuss this finding in the context of the hypothesis that Parkinson's disease may stem from clinically silent damage to the substantia nigra, followed by slow attrition of neurones in this region because of its particular vulnerability to cell loss as a normal consequence of ageing.

Two of our subjects are the father and sister, respectively, of a case of parkinsonism induced by MPTP, 'patient 6' reported by Langston and Ballard². They were injecting the same 'synthetic heroin' as 'patient 6', but in lower doses. Both had acute transient symptoms in 1982 at the time of exposure, typical of all subjects receiving MPTP: a burning sensation at the site of injection, blurring of vision, a metallic taste in the mouth and jerking of the limbs³. The other two subjects were also supplied with synthetic heroin in 1982. They both had a burning sensation at the site of injection; one had a metallic taste in his mouth and the other experienced blurring of vision. Computerized tomography or magnetic resonance imaging were normal in all

four cases. Clinical examinations were performed, together with objective tests developed in our laboratory⁴, for tremor, rigidity, velocity and accuracy of movement. Two of the MPTP subjects were entirely normal and two had minimal disturbances which were insufficient for any pathology to be diagnosed. Two of the subjects had been off narcotics for 3 months; the other two subjects were receiving methadone, but none had been taken for over 5 h before the PET scan.

For comparative purposes we also studied seven normal subjects and six patients with spontaneous Parkinson's disease (Hoehn and Yahr stages I to III), none of whom had been exposed to MPTP. Blood biochemistry, including tests of liver function, were normal in all subjects. The ages of the subjects exposed to MPTP were 26, 29, 36 and 48 yr, and the normal subjects were aged 28, 43, 52, 54, 67, 69 and 73 yr; the parkinsonian age range was similar to that of the normal subjects.

Positron-emitting ¹⁸F was produced with a CP-42 cyclotron as ¹⁸F-F₂ by the ^{nat}Ne(p, x)¹⁸F reaction. The ¹⁸F-F₂ gas was converted to acetyl hypofluorite and then incorporated into a dopa derivative using the procedure described by Adam *et al.*⁵. Using this method, we were able to prepare ¹⁸F-6-fluoro-L-dopa (6-FD) routinely with 95% radiochemical purity; the product was 99% in the L-form and, after HPLC purification, was fluorinated >96% in the 6-position. The subjects received 3-4 mCi (equivalent to ~2 mg) of 6-FD 1 h after the oral administration of 100 mg carbidopa, given to prevent peripheral decarboxylation.

Twelve sequential scans, each lasting 10 min, were performed immediately after isotope administration using the UBC-TRIUMF PET VI tomograph, whose characteristics have been described previously^{6,7}. This scanner allows the simultaneous acquisition of 7 slices of emission data, with a slice thickness and centre-to-centre separation of 14 mm, and in-slice resolution of 8 mm. By taking alternate scans with the subject shifted in the axial direction by 7 mm (half a slice thickness), 14 interleaving slices were obtained with an effective centre-to-centre spacing of 7 mm. Attenuation correction was accomplished by performing two transmission scans using a ⁶⁸Ge-containing ring source, in positions identical to the emission scans. Major head movements were prevented by a moulded thermoplastic face mask. All subjects were positioned with the aid of a vertical laser line, such that the lowest slice was at the level of the inferior orbito-meatal line. Reconstructed images were displayed in colour-coded form on a 100 × 100-pixel matrix. To minimize inter-subject variability, all striatal data analysed were from the two adjacent slices in which striatal radioactivity was the highest, selected from the total set of 14 transverse images from each subject.

Comparable axial scans from a normal subject, a parkinsonian patient and a MPTP-exposed subject without clinical deficits are shown in Figs 1 and 2. Figure 3 shows the accumulation of

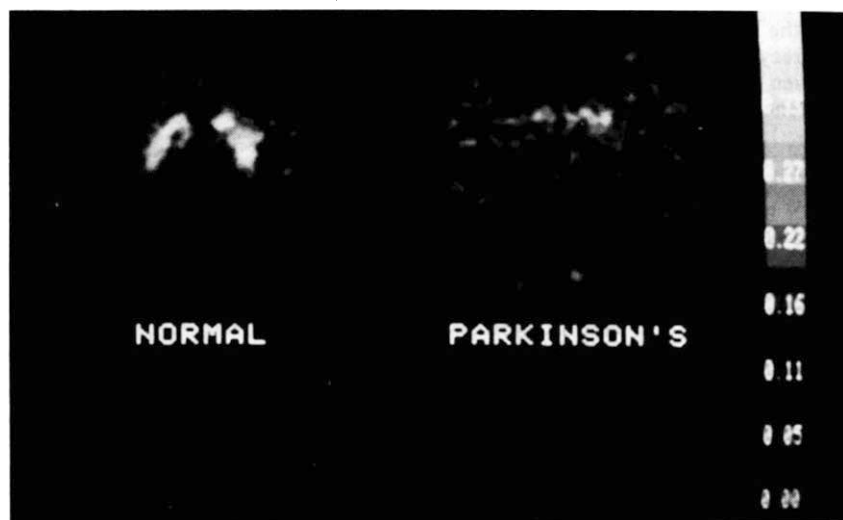
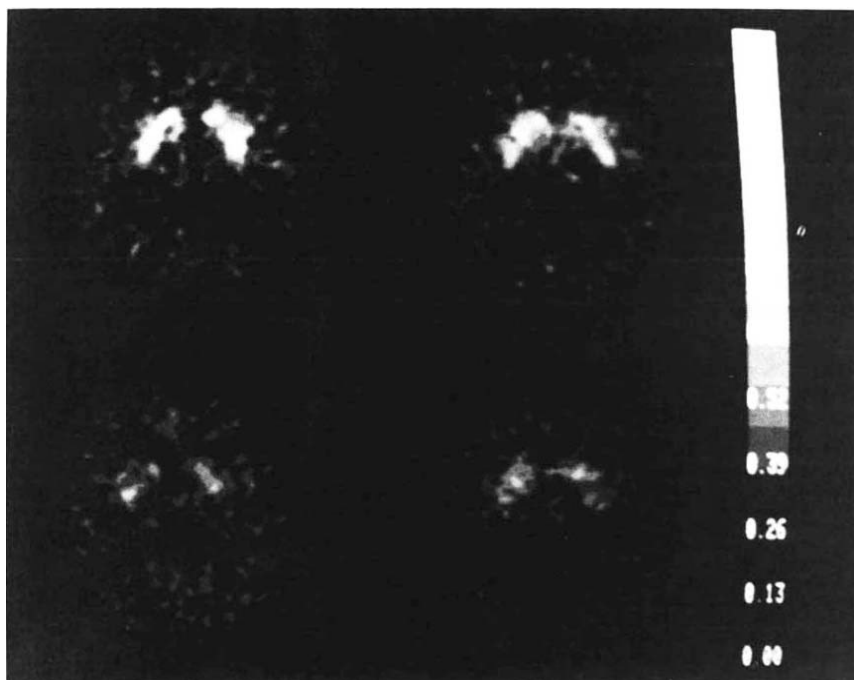


Fig. 1 PET scans, axial view through the same level from a normal subject (left) and a patient with spontaneous Parkinson's disease, Hoehn and Yahr stage I (right). Activity accumulated over a period of 60-120 min following intravenous fluorodopa. The scale on the right represents radioactivity in linear arbitrary units, increasing from black to white; reconstructed to give similar background (cerebellar) activity. In the parkinsonian patient, activity is moderately reduced in the caudate and severely decreased in the putamen, as others have reported¹⁶.

Fig. 2 Upper: fluorodopa PET scans from another normal subject, similar to the left scan in Fig. 1, but showing two adjacent axial cuts through the striatum. Lower: fluorodopa PET scans from a subject without clinical deficits but exposed to MPTP, taken at the same levels as in the normal scans above. The activity is intermediate between normality and Parkinson's disease.



striatal activity over the period 60–120 min after administration of 6-FD, divided by corresponding cerebellar activity. These ratios were obtained by adding regional data from the individual scans performed in both positions during the second hour, and averaging the values. The same procedure was followed for all subjects. Values over the first hour were not included because these were thought to reflect tracer delivery and early tissue distributions of 6-FD and metabolites (analogous to the distribution of ^3H -dopa and metabolites shown by Horne *et al.*⁸). For assessment of the results, a logarithmic transformation was applied to the data to stabilize the variances. Analysis of variance revealed significant differences between the three means ($P < 0.001$; $F(2, 15) = 12.58$). The Newman-Keuls method was used to compare the effects in pairs; the normal mean was significantly higher than the MPTP and parkinsonian means ($P < 0.05$). On the basis of the hypothesis proposed by Calne and Langston⁹, a one-sided *t*-test was performed. Although dependent data were used for further analysis, computation revealed a significant reduction in the MPTP subjects compared with normals ($P = 0.03$, $t = 2.20$; 9 d.f.); similarly, the patients with Parkinson's disease had significantly lower values than the MPTP subjects ($P = 0.03$, $t = 2.16$; 9 d.f.). Thus, the ratio of striatal to cerebellar activity in the brain of these MPTP-exposed subjects represents an intermediate stage between normality and Parkinson's disease.

The technique for analysis of 6-FD uptake is, at present, only semi-quantitative. More precise analysis of the observations must await the development of mathematical modelling of 6-FD kinetics. Preliminary work in this area has been reported^{10,11}. However, because we found that the plasma time-activity curve, the relative plasma concentrations of labelled 6-FD and its metabolites, and the cerebellar activity/injected dose ratio were all similar in the MPTP-exposed subjects and controls, we suggest that a real disparity exists between the striatal activity of the two groups. Although we cannot distinguish between uptake, storage or degradation of fluorodopa or its metabolites, we can infer that the reduction in striatal radioactivity in the MPTP subjects stems from a disturbance of dopaminergic nerve endings because this is the only explanation compatible with current physiological, biochemical and pharmacological concepts of basal ganglia function. Dopaminergic endings in the striatum all originate in the nigrostriatal pathway, which is the focal point of neuronal degeneration in Parkinson's disease.

The action of opiates on cerebral dopamine stores is not well documented in human subjects. There is one case report of

transient parkinsonism in a patient receiving a drug regimen that included pethidine¹²; however, induction of a brief parkinsonian syndrome has been reported as a nonspecific reaction to many drugs. There is no recognized association between narcotics and parkinsonism, except through contamination with MPTP. Opiate derivatives do not exacerbate the symptoms or signs in patients with parkinsonism; furthermore, opiate antagonists have no therapeutic action in Parkinson's disease¹³. Most of the relevant observations therefore militate against the possibility that opiates would impede striatal accumulation of 6-FD or its metabolites. Indeed, animal experiments indicate that opiates increase the rate of dopamine synthesis¹⁴.

The conclusion that subclinical damage to the nigrostriatal pathway can occur in human subjects has implications for the aetiology of Parkinson's disease. It has been suggested that the disease may result from an environmental event (or series of events) preceding the onset of symptoms by many years⁹, and that the subsequent normal age-related loss of dopaminergic neurones erodes the compromised substantia nigra to result in the emergence and progression of clinical deficits. The findings reported here are consistent with this hypothesis. It is relevant

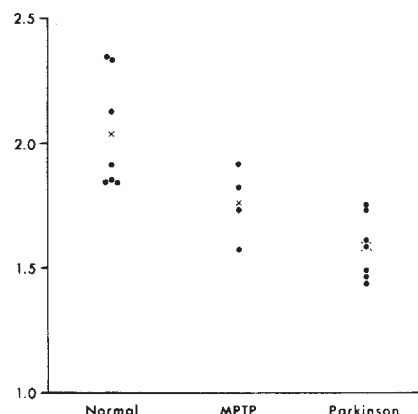


Fig. 3 Striatal activity determined from fluorodopa PET scans. Each point represents a value for bilateral striatal (caudate plus putamen) activity divided by background (cerebellar) activity. The cross shows the mean for each group. Results for the MPTP subjects are significantly less than normal, but more than those for patients with Parkinson's disease.

that one of us (J.L.) has documented progressive micrographia in a patient with parkinsonism following exposure to MPTP¹⁵. Patients with post-encephalitic parkinsonism represent an interesting analogy with subjects exposed to MPTP. In both cases parkinsonism follows a reasonably well-defined environmental event. One of us (D.B.C.) has observed a group of 50 parkinsonian patients surviving the pandemic of encephalitis lethargica; after remaining relatively stable over many years, most patients decompensated from the age of about 60 yr, when the neurological deficits progressed. This deterioration is compatible with an aetiological hypothesis of environmental damage superimposed on dopaminergic attrition caused by ageing.

Finally, we emphasize that the hypothesis of environmental hazard plus ageing allows considerable flexibility. A widespread causal agent, whether toxic or infective, may have diverse actions in a variety of subjects because of individual variations in metabolism or immunity that may be necessary to achieve activation or inactivation. Such differences could explain the difficulty in finding consistent environmental risk factors, and also the considerable disparity in the rate of progress of Parkinson's disease in different patients.

This work was supported by the Dystonia Medical Research Foundation and the MRC of Canada. We thank Edwin Mak,

Pat Porter and Susan Calne for their assistance, Dr John Grierson and Selma Jivan for their help in preparing the scanning agent, and the members of the UBC/TRIUMF PET team.

Received 7 May; accepted 1 August 1985.

1. Garnett, E. S., Firnau, G. & Nahmias, C. *Nature* **305**, 137-138 (1983).
2. Langston, J. W. & Ballard, P. *Can. J. Neurol. Sci.* **11**, 160-166 (1985).
3. Langston, J. W. *Life Sci.* **36**, 201-206 (1985).
4. Teravainen, H. T. & Calne, D. B. in *Parkinson's Disease—Current Progress, Problems and Management* (eds Rinne, U.K., Klinger, M. & Stamm, G.) 145-165 (Elsevier/North-Holland, Amsterdam, 1980).
5. Adam, M. J., Abeysekera, B., Ruth, T. J., Grierson, J. R. & Pate, B. D. *J. Nucl. Med.* **26**(5), P125 (1985).
6. Ter-Pogossian, M. M., Ficke, D. C., Hood, J. T., Yamamoto, M. & Mullani, N. A. *J. Computer Assisted Tomography* **6**, 125-133 (1982).
7. Evans, B. et al. *IEEE Trans. Nucl. Sci.* **NS-30**, 707-710 (1983).
8. Horne, M. K., Chang, C. H. & Wooten, G. F. *Pharmacology* **28**, 12-26 (1984).
9. Calne, D. B. & Langston, J. W. *Lancet* **ii**, 1457-1459 (1983).
10. Garnett, E. S., Firnau, G., Nahmias, C., Sood, S. & Belbeck, L. *Am. J. Physiol.* **238**, R318-R327 (1980).
11. Martin, W. R. W., Boyes, B. E., Leenders, K. L. & Patlak, C. S. *J. Cerebral Blood Flow Metab.* (in the press).
12. Lieberman, A. N. & Goldstein, M. *New Engl. J. Med.* **312**, 509 (1985).
13. Nutt, J. G., Rosin, A. J., Eisler, T., Calne, D. B. & Chase, T. N. *Archs Neurol.* **35**, 827-830 (1978).
14. Kushinsky, K. & Hornykiewicz, O. *Eur. J. Pharmacol.* **Vol 74**:1-50 (1974).
15. Langston, J. W. in *Recent Advances in Parkinson's Disease* (eds Fahn, S., Marsden, C. D., Jenner, P. & Teychenne, P.) (Raven, New York, in the press).
16. Garnett, E. S., Nahmias, C. & Firnau, G. *Can. J. Neurol. Sci.* **11**, 174-179 (1984).

Co-localization of corticotropin-releasing factor and vasopressin in median eminence neurosecretory vesicles

Mark H. Whitnall*, Éva Mezey† & Harold Gainer*

* Laboratory of Neurochemistry and Neuroimmunology, National Institute of Child Health and Human Development, National Institutes of Health, Bethesda, Maryland 20892, USA

† Laboratory of Cell Biology, National Institute of Mental Health, Bethesda, Maryland 20892, USA

Vasopressin (VP) potentiates the effect of corticotropin-releasing factor (CRF) on the secretion of adrenocorticotrophic hormone (ACTH) from anterior pituitary cells *in vitro*¹⁻³, and both CRF⁴⁻⁷ and VP^{6,8} have been found in portal blood. These data support the hypothesis that VP acts synergistically with CRF to cause the secretion of ACTH *in vivo*^{1-4,8,9} but the origin of the CRF and VP, and the physiology of their release, have not been precisely defined. Parvocellular cell bodies in the paraventricular nucleus (PVN) which project to the external zone of the median eminence can be stained for both CRF and VP after adrenalectomy¹⁰⁻¹², and there is light microscopic immunocytochemical evidence that neuropeptide Y (NP) may be located within some of the CRF-containing axons¹³. Electron microscopic immunocytochemical studies have demonstrated the presence of CRF¹⁴⁻¹⁶, VP^{17,18} and its 'carrier' protein, VP-associated neuropeptide Y (NP-VP)^{17,18} in 100-nm neurosecretory vesicles (NSVs) in axons terminating near the portal capillary plexus in the external zone of the median eminence. If these peptides are extensively co-localized in the same NSVs in the median eminence, then coordinate secretion of CRF and VP *in vivo* is obligatory, at least in some physiological circumstances. We demonstrate in this report, using post-embedding electron microscopic immunocytochemistry on serial ultrathin sections, that CRF, VP and NP-VP are contained not only in the same axons and terminals, but in the same 100-nm NSVs in the median eminence of both normal and adrenalectomized rats. In addition, in the normal rat median eminence 44% of the CRF-positive axons and terminals stained strongly for VP and NP-VP, whereas in the adrenalectomized rat virtually all the CRF-positive structures in the median eminence showed strong staining for VP and NP-VP, indicating a transformation of one subpopulation of CRF-positive axons and terminals by adrenalectomy.

Adult male rats were decapitated and their brains immersed in 4% glutaraldehyde, 0.2% picric acid, 0.1M sodium cacodylate, pH 6 for 2 h. In some cases, bilateral adrenalectomy was performed under ether anaesthesia 3 days before fixation. Horizontal ventral surface slabs containing the median eminence were then dissected out, and some blocks were post-fixed in 1% osmium tetroxide for 2 h. Tissues were embedded in LR White resin¹⁹, and serial sections were picked up so that staining could be done on the back of one section and the front of the next section. Immunoperoxidase staining of the sections was performed as previously described²⁰. Sections of osmicated tissue were pre-incubated in 0.5% sodium metaperiodate for 10 min to restore antigenicity²¹. No differences in staining patterns for CRF and NP-VP were seen between the osmicated and unsmicated tissues, but the parvocellular axons did not stain for VP after osmication.

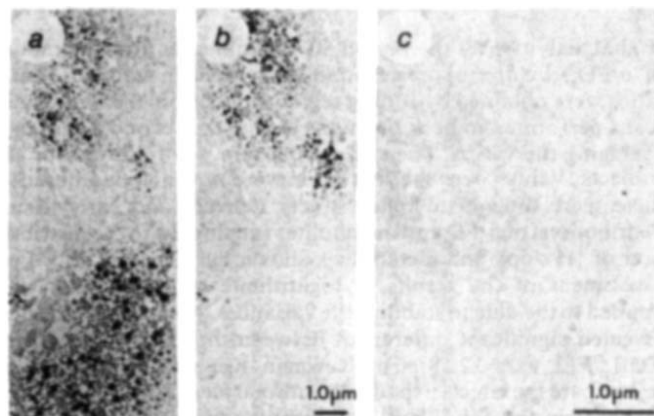


Fig. 1 Co-localization of CRF and NP-VP in parvocellular axons in the external zone of the median eminence. *a, b*, Serial sections of osmicated tissue from the median eminence of an adrenalectomized rat stained for NP-VP (*a*) or CRF (*b*). The parvocellular axons (top) contain 100-nm NSVs which stained for both CRF and NP-VP. In contrast, the magnocellular axon (bottom) contains 160-nm NSVs which stained for NP-VP but not for CRF. *c*, An absorption control for CRF staining. Parvocellular axons were exposed to CRF antiserum absorbed with 5×10^{-5} M synthetic ovine CRF.

The CRF antibody used in these studies was a rabbit antiserum against rat CRF, and its immunocytochemical specificity is described elsewhere²². This CRF antiserum was used at 1:800, and stained 100-nm NSVs in the external zone of the median eminence (Fig. 1*b*); magnocellular axons (containing 160-nm

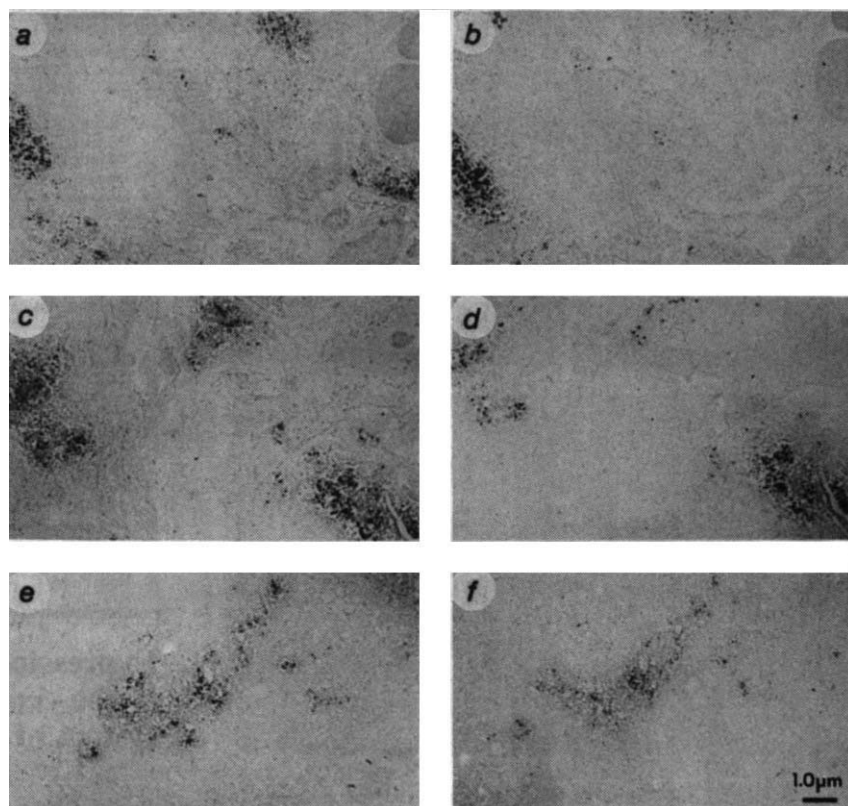


Fig. 2 *a, b*, Serial sections of osmicated tissue from the external zone of a normal rat; *c, d* and *e* and *f* are serial sections of unosmicated tissue from an adrenalectomized rat. Sections were stained for CRF (*a, c, e*), NP-VP (*b, d*) or VP (*f*). Many CRF-positive axons stain only faintly for NP-VP in the normal rat, but almost all CRF-positive axons are clearly stained for NP-VP in the adrenalectomized rat. A similar correspondence is seen on the serial sections of unosmicated tissue from an adrenalectomized rat stained for CRF (*e*) or VP (*f*).

NSVs) in the internal zone did not stain for CRF in these conditions (Fig. 1*b*), but did stain at higher antibody concentrations (not shown). Other studies have detected CRF in the oxytocin (OT) magnocellular neurones which project axons through the internal zone of the median eminence²²⁻²⁴, in the 160-nm NSVs²⁴. Staining of the parvocellular axons was completely blocked by the addition of 5×10^{-5} M synthetic ovine CRF to the antiserum (Fig. 1*c*). An antiserum against VP was used at 1:500. This antiserum is specific for VP in radioimmunoassays (M. Altstein and G. Valiquette, personal communications), does not stain the OT-containing axons in the internal zone of the median eminence, as shown by serial section studies (unpublished data), and staining could be absorbed out with VP-Sepharose beads (not shown). As VP and NP-VP are synthesized on the same prohormone²⁵ and are always colocalized in the same NSVs²⁶, immunocytochemical evidence for the presence of NP-VP also indicates the presence of VP. To stain for NP we used a monoclonal antibody, PS 45, that is highly specific for NP-VP in liquid-phase assays, but can cross-react with OT-associated NP (NP-OT) in immunocytochemistry^{27,28}. However, PS 45 stained only NP-VP in the parvocellular axons in the external zone of the median eminence, as an equally specific and potent antibody (PS 36) against NP-OT, which does not cross-react with NP-VP in immunocytochemical tests^{27,28}, could not stain these axons in several light and electron microscopic studies. This finding is consistent with previous observations that OT fibres are extremely scarce in the external zone of the median eminence, and the few that exist do not respond to adrenalectomy in the same way as VP fibres²⁹. Hence these scarce OT fibres are probably stray magnocellular axons from the hypothalamo-neurohypophyseal system, and not part of a parvocellular CRF/OT projection to the portal capillaries.

Figures 1 and 2 illustrate electron microscopic immunocytochemistry of serial sections stained for either CRF or NP-VP (or VP) at low magnification. At these magnifications it is possible to localize, in some sections, both 100-nm vesicle-containing and 160-nm vesicle-containing axons in the median eminence, and to show that only the parvocellular axons (containing 100-

nm vesicles) show coexistence of CRF and NP-VP (Fig. 1). In both normal and adrenalectomized rats, all NP-VP- or VP-positive parvocellular axons and terminals examined in the median eminence were also stained for CRF (Figs 1, 2). However, in the normal rat, not all CRF-positive axons stained for NP-VP or VP. This is illustrated in Fig. 3, where staining of serial sections revealed that in some axons, identical 100-nm NSVs contained strong immunoreactivity for both CRF and NP-VP, whereas in others, CRF-positive axons stained faintly or not at all for NP-VP. In every pair of serial sections analysed, from either normal or adrenalectomized rats, the same NSVs could easily be identified by their identical patterns on the two sections (Fig. 3, arrows). Similar results were obtained using the VP antiserum (not shown). Although this indication of two CRF-positive parvocellular axon subpopulations (with and without VP coexistence) was characteristic of the normal rat median eminence, many more of the CRF-positive axons in the adrenalectomized rat showed strong co-staining for NP-VP or VP. In a preliminary quantitation of this phenomenon, 82 of 187 CRF-positive axons and terminals (44%) in the external zone of the median eminence were found to also stain intensely for NP-VP in normal, unoperated rats, whereas 105 (56%) were only faintly stained or unstained for NP-VP. However, 3 days after bilateral adrenalectomy, 141 of 157 CRF-positive axons (or 90%) in the median eminence were strongly stained for NP-VP or VP. This increase in NP-VP and VP staining in the CRF-positive parvocellular axons after adrenalectomy is consistent with previous observations^{18,29-32} of increases in immunocytochemical staining for VP and NP-VP in the external zone, the appearance of VP staining¹⁰⁻¹² and VP messenger RNA³³ in CRF-containing cell bodies in the paraventricular nucleus, and the increase in axonal transport of newly synthesized VP and NP-VP to the median eminence³⁴ after adrenalectomy.

The findings described here lead to two unequivocal conclusions: (1) CRF, VP and NP-VP are not only coexistent in the same axons in the external zone of the median eminence, but also in the same 100-nm NSVs in those axons. This would imply co-secretion of these peptides by those axon terminals.

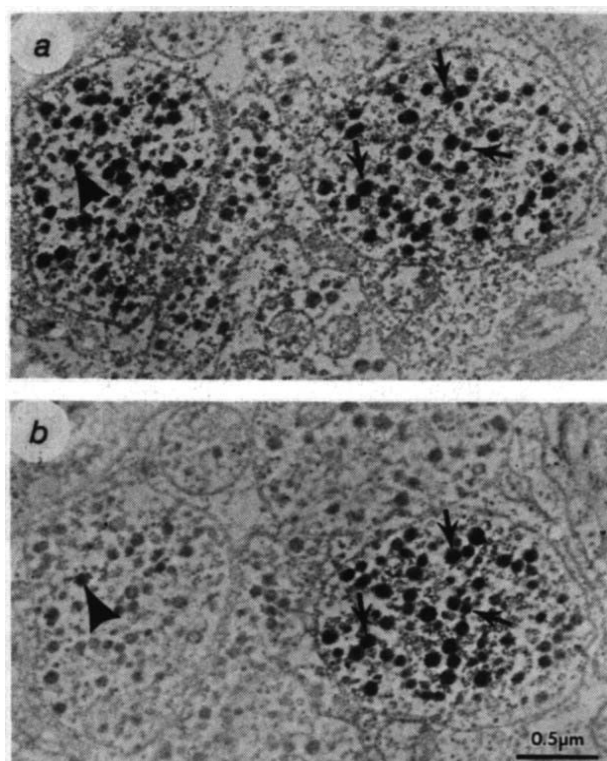


Fig. 3 Serial LR White sections of osmicated tissue from the external zone of the median eminence of a normal rat, stained for CRF (a) or NP-VP (b). The axonal swelling on the right stained intensely for both CRF and NP-VP. Examples of identical NSVs labelled on both sections are indicated by arrows. The swelling on the left stained intensely for CRF, but only a few NSVs are faintly stained for NP-VP, one of which is indicated by an arrowhead.

(2) At least two subpopulations of CRF-positive axons and terminals exist in the external zone of the median eminence in the normal rat—that is, one (~44%) in which VP is found to abundantly coexist (by electron microscopic immunocytochemical criteria), and another (~56%) in which VP and NP-VP are undetectable, or barely detectable, by the same criteria. These results suggest that modulation of the CRF/VP ratio in portal blood could be accomplished by differential activation of subpopulations of parvocellular axons in normal animals. After adrenalectomy, however, this potential flexibility is apparently overridden by the strong stimulus for ACTH release, so that a relatively homogeneous population of parvocellular axons emerges containing high levels of both CRF and VP. The physiological significance of these findings must await further information about the concentrations of CRF and VP in portal blood in various physiological conditions, and their effects on the separate receptors for CRF and VP in the adeno-hypophysial corticotrophs³⁵. In any case, these data on the co-packaging of CRF and VP provide cell biological evidence in support of the hypothesis that VP plays an important role in the central regulation of ACTH release.

We thank Dr W. Vale for CRF antiserum, Drs E. A. Zimmerman and G. Valiquette for VP antiserum, Dr F. Antoni for helpful discussions, Drs G. Newman and B. Jasani for introducing us to the use of LR White for ultrastructural immunocytochemistry, and S. Key for technical assistance.

Received 2 April; accepted 26 July 1985.

1. Gillies, G. E., Linton, E. A. & Lowry, P. J. *Nature* **299**, 355–357 (1982).
2. Turkelson, C. M. *et al. Peptides* **3**, 111–113 (1982).
3. Vale, W. *et al. Endocrinology* **113**, 1121–1131 (1983).
4. Vale, W. W., Rivier, C., Spiess, J. & Rivier, J. in *Brain Peptides* (eds Krieger, D., Martin, J. B. & Brownstein, M. J.) 961–974 (Wiley, New York, 1983).
5. Gibbs, D. M. & Vale, W. *Endocrinology* **111**, 1418–1420 (1982).
6. Gibbs, D. M. & Vale, W. *Brain Res.* **280**, 176–179 (1983).
7. Plotsky, P. M., Bruhn, T. O. & Vale, W. *Endocrinology* **115**, 1639–1641 (1984).

8. Zimmerman, E. A. *et al. Ann. N.Y. Acad. Sci.* **297**, 405–419 (1977).
9. Yates, F. E. *et al. Endocrinology* **88**, 2–15 (1971).
10. Tramu, G., Croix, C. & Pillez, A. *Neuroendocrinology* **37**, 467–469 (1983).
11. Kiss, J. Z., Mezey, E. & Skirboll, L. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1854–1858 (1984).
12. Sawchenko, P. E., Swanson, L. W. & Vale, W. W. *Proc. natn. Acad. Sci. U.S.A.* **81**, 1883–1887 (1984).
13. Tramu, G. & Pillez, A. *C.r. hebdom. Séanc. Acad. Sci., Paris Ser. III* **294**, 107–114 (1982).
14. Pelletier, G. *et al. Neuroendocrinology* **35**, 402–404 (1982).
15. Leranthe, C., Antoni, F. A. & Palkovits, M. *Regul. Peptides* **6**, 179–188 (1983).
16. Liposits, Z. *et al. Cell Tissue Res.* **229**, 191–196 (1983).
17. Silverman, A. J. & Zimmerman, E. A. *Cell Tissue Res.* **159**, 291–301 (1975).
18. Dube, D., Leclerc, R. & Pelletier, G. *Am. J. Anat.* **147**, 103–108 (1976).
19. Newman, G. R., Jasani, B. & Williams, E. D. *Histochem. J.* **15**, 543–555 (1983).
20. Whitnall, M. H., Castel, M., Key, S. & Gainer, H. *Peptides* **6**, 241–247 (1985).
21. Bendayan, M. & Zollinger, M. *J. Histochem. Cytochem.* **31**, 101–109 (1983).
22. Sawchenko, P. E., Swanson, L. W. & Vale, W. W. *J. Neurosci.* **4**, 1118–1129 (1984).
23. Burlet, A., Tonon, M.-C., Tankosic, P., Coy, D. & Vaudry, H. *Neuroendocrinology* **37**, 64–72 (1983).
24. Dreyfuss, F., Burlet, A., Tonon, M.-C. & Vaudry, H. *Neuroendocrinology* **39**, 284–287 (1984).
25. Land, H., Schutz, G., Schmale, H. & Richter, D. *Nature* **295**, 299–303 (1982).
26. Castel, M., Gainer, H. & Dellman, H.-D. *Int. Rev. Cytol.* **88**, 303–459 (1984).
27. Ben-Barak, Y., Russell, J. T., Whitnall, M. H., Ozato, K. & Gainer, H. *J. Neurosci.* **5**, 81–97 (1985).
28. Whitnall, M. H., Key, S., Ben-Barak, Y., Ozato, K. & Gainer, H. *J. Neurosci.* **5**, 98–109 (1985).
29. Stillman, M. A. *et al. Endocrinology* **101**, 42–49 (1977).
30. Chateau, M., Burlet, A. & Marchetti, J. *J. Physiol., Paris* **68**, 10B–11B (1974).
31. Vandesande, F., DeMey, J. & Dierckx, K. *Cell Tissue Res.* **151**, 187–200 (1974).
32. Watkins, W. B., Schwabedal, P. & Bock, R. *Cell Tissue Res.* **152**, 411–421 (1974).
33. Wolfson, B., Manning, R. W., Davis, L. G., Arentzen, R. & Baldino, F. *Nature* **315**, 59–61 (1985).
34. Russell, J. T., Brownstein, M. J. & Gainer, H. *Brain Res.* **201**, 227–234 (1980).
35. Holmes, M. C., Antoni, F. A. & Szentendrei, T. *Neuroendocrinology* **39**, 162–169 (1984).

Expression of human HPRT in the central nervous system of transgenic mice

J. Timothy Stout*, Howard Y. Chen†, John Brennan*, C. Thomas Caskey* & Ralph L. Brinster†

* Departments of Medicine and Cell Biology, M-804 DeBakey Bldg, Baylor College of Medicine, One Baylor Plaza, Houston, Texas 77030, USA

† Laboratory of Reproductive Physiology, School of Veterinary Medicine, University of Pennsylvania, Philadelphia, Pennsylvania 19104, USA

Severe deficiency of hypoxanthine phosphoribosyltransferase (HPRT) in man results in the Lesch–Nyhan syndrome, an X-linked neurological disorder characterized by mental retardation, choreoathetosis and a compulsive tendency towards self-mutilation¹. Although the HPRT gene is normally constitutively expressed in all tissues at low levels, expression is elevated approximately fourfold in several regions of the central nervous system, particularly in the basal ganglia^{2,3}. The relationships between HPRT deficiency, tissue-specific alterations of nucleotide metabolism and the neuropathology of the Lesch–Nyhan syndrome remain unclear. Here we have microinjected recombinant molecules containing human HPRT (hHPRT) complementary DNA, the mouse metallothionein-I (MT-I) promoter and the 3'-untranslated portion of the human growth hormone (hGH) gene into mouse embryos to produce transgenic animals that express hHPRT on induction by cadmium. The hHPRT cDNA in these experiments contained 88 base pairs (bp) of 5'-untranslated and 190 bp of 3'-untranslated sequences, and the full-length coding sequence. We studied the *in vivo* expression of this MT-hHPRT fusion gene and observed preferential hHPRT expression in tissues of the central nervous system (CNS). This study suggests that sequences within the hHPRT transcript (cDNA) influence CNS expression via increased synthesis or stability of messenger RNA.

Of 389 mouse embryos injected with ~900 copies of the linearized plasmid pHPT-50 (see Fig. 1 for a description of the plasmid), 241 were selected for transfer to pseudopregnant females. Seventeen of the 45 offspring contained from one to five stably integrated copies of the fusion gene. Southern analysis of DNAs from the four mice that expressed this fusion gene indicated that the genes had integrated into a single chromosomal location¹³. The junctional fragments suggested that each

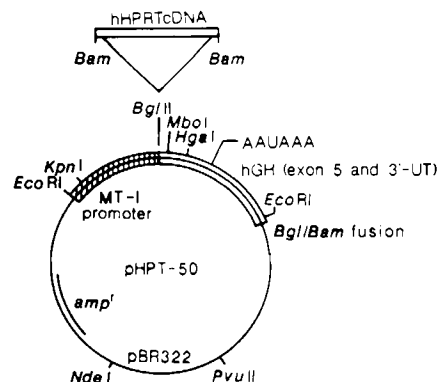


Fig. 1 Human HPRT cDNA was cloned into the *Bgl*II site of a plasmid containing ~770 bp of sequence 5' to the mouse MT-I gene and 650 bp of 3' flanking sequence from the human growth hormone gene (MT-I/hGH expression vector provided by R. Palmiter). We injected 2 μ l containing 900 linearized plasmids into the male pronucleus of mouse embryos as described elsewhere¹¹. Embryos were transferred to pseudopregnant surrogate mothers, and, after weaning, tail DNA from the resulting pups was used to detect carriers of the fusion gene¹².

Table 1 HPRT activity in transgenic mice and their progeny

Animal	HPRT species	pHPT-50 copy no.	Tissue					
			Brain	Heart	Kidney	Liver	Lung	Spleen
15-1	Human	1	8.37	1.92	0.13	0.78	0.57	0.63
(F ₀)	Mouse		0.50	0.20	2.82	2.34	2.27	5.57
15-1-1	Human	1	11.52	2.54	0.30	1.12	0.68	0.69
(F ₁)	Mouse		1.25	0.56	3.82	3.42	3.46	7.89
18-5	Human	3	10.91	0	4.85	0	0	0
(F ₀)	Mouse		3.25	2.17	2.00	2.11	2.51	4.37
18-5-1	Human	3	9.30	0	4.31	0	0	0
(F ₁)	Mouse		3.73	2.09	2.02	2.07	2.34	4.20
18-9	Human	2	10.48	3.40	2.36	0.99	1.40	4.47
(F ₀)	Mouse		6.65	2.01	3.51	2.30	2.49	5.29
23-4	Human	1	17.20	0	0	7.93	0	0
(F ₀)	Mouse		3.50	1.37	1.41	0.71	1.31	1.99
Normal mouse control	Mouse	0	4.8	1.2	1.4	1.2	1.4	2.4

Activity expressed as nmol IMP formed per min per mg of protein. Normal mouse control values are mean values for eight non-transgenic mice.

mouse had a different site of gene integration (data not shown).

Human HPRT expression in brain, heart, kidney, liver, lung and spleen tissues was determined by enzyme and mRNA quantitation. The proportion of human and murine HPRT enzyme synthesized was determined by enzyme assay following their separation by isoelectric focusing (Fig. 2). The concentrations of hHPRT mRNA were estimated by 'slot-blot' analysis (Fig. 3) using a probe from the 3'-untranslated region of the fusion gene.

Four of eight mice examined in detail expressed the human gene following Cd²⁺ induction⁴. Two mice (15-1 and 18-9) expressed the gene in all tissues; a third mouse (23-4) showed expression in brain and liver tissues only; and the fourth (18-5) expressed the gene only in brain and kidney tissues. In each case, expression of hHPRT was highest in brain tissues (Table 1). Mouse 15-1 showed a concentration of hHPRT that was 16-fold higher than the concentration of endogenous mouse enzyme. Mouse 23-4 exhibited a greater than fourfold increase of total CNS HPRT activity compared with non-transgenic control mice, without phenotypic effect. In addition, in each transgenic mouse, expression of the fusion gene was consistently elevated (up to 64-fold) in brain tissues compared with expression in other tissues. The level of hHPRT expression did not correlate with the copy number of the MT-hHPRT gene. There was an approximate correlation between the concentra-

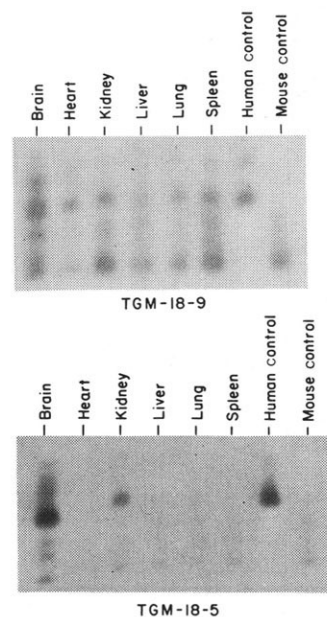


Fig. 2 Representative isoelectric focusing (IEF) gels of transgenic mice.

Methods. The mouse organs shown (0.1 g of each) were homogenized in 500 μ l of 20 mM Tris-HCl (pH 7.2) in a Dounce homogenizer and subjected to five freeze-thaw cycles to lyse cells. Cellular debris was removed by centrifugation and 5 or 10 μ l of the cytosolic supernatant was loaded onto a non-denaturing IEF gel (10 cm $\times$ 21 cm $\times$ 0.5 mm) and focused at 700 V for 12.5 h (pH 5-7/21 cm). The gel was overlaid with a reaction cocktail containing 100 mM Tris-HCl (pH 7.4), 10 mM MgCl₂, 1 μ Ci ¹⁴C-hypoxanthine (50 mCi mmol⁻¹), 10 mM phosphoribosylpyrophosphate and 4.8 mg ml⁻¹ bovine serum albumin (1 ml cocktail per 45 cm² of gel) and incubated at 37°C for 1 h^{14,15}. A sheet of polyethyleneimine-impregnated cellulose was placed on the gel and ¹⁴C-IMP was allowed to adsorb for 15 min at room temperature. The cellulose sheet was washed twice with H₂O at room temperature and autoradiographed. The fractional representation of human and mouse enzymes was determined by cutting out radioactive spots corresponding to each enzyme, followed by quantitation by liquid scintillation counting. Total activity was determined by direct assay¹⁶.

tion of hHPRT mRNA in the CNS and CNS hHPRT-specific activity.

Offspring from two of the four mice were also examined for gene inheritance and expression. Mouse 15-1, a male that expressed the human gene in all tissues, sired eight pups, five of which inherited the human gene. In one of the male offspring (15-1-1) the hHPRT gene was expressed in all tissues in a manner quantitatively similar to the parent. Mouse 18-5, a male that showed expression in brain and kidney tissues, transmitted the gene to 6 of 11 offspring; one of the male offspring, 18-5-1, expressed hHPRT only in brain and kidney tissues at levels comparable to those of its parent. Direct male-to-male transmission of hHPRT tissue-specific expression supports an autosomal dominant mendelian pattern of inheritance of the human gene.

The responsiveness of the fusion gene to heavy-metal induction was examined by Northern analysis of RNAs isolated from induced or uninduced transgenic siblings. Figure 4 demonstrates the inducible nature of the gene. Cadmium treatment of mouse 15-1-2, a female offspring of mouse 15-1, resulted in increased synthesis of HPRT mRNA in all tissues compared with the levels observed in an untreated female sibling, mouse 15-1-3. Offspring of mouse 18-5 demonstrated a similar tissue-specific pattern of inducibility in kidney and brain tissues. While human and mouse HPRT transcripts could not be separated by size,

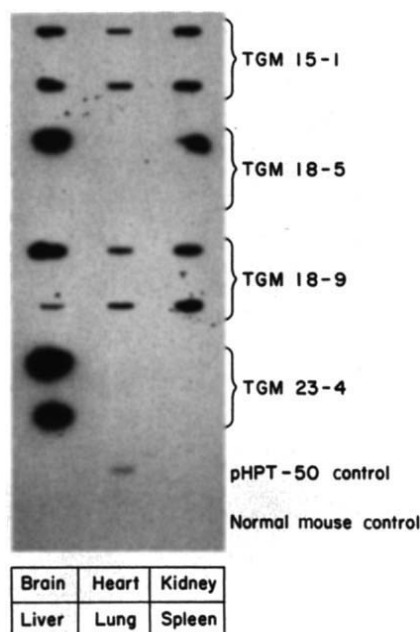


Fig. 3 Slot-blot analysis of RNAs from transgenic mice. RNA was isolated from fresh tissues of each mouse¹⁷. Total RNA (20 µg) was immobilized on nitrocellulose and hybridized to nick-translated *HgaI-EcoRI* fragment from pHPT-50 in 50% formamide, 5 × SSC, 0.1% SDS, and 1 × Denhardt's at 65 °C for 12 h^{17,18}. RNA from HPRT⁺ Chinese hamster cells transfected with pHPT-50 was used as a positive control and RNA from a non-transgenic mouse brain was used as a negative control. The key at the bottom indicates the tissue source of RNA for each mouse.

heavy-metal inducibility is indicative of a functional metallothionein promoter, active in generating hHPRT transcripts.

The mechanism underlying the unexpected elevation of MT-hHPRT gene expression in the CNS of transgenic mice remains unclear. Durnam and Palmiter have demonstrated that induction of genes under the control of the mouse metallothionein promoter involves an increase in the rate of transcription of those genes and that this effect is primarily limited to the liver and kidney⁵. Cd²⁺ induction of these genes in brain tissue resulted in no more than a 1.4-fold increase in gene transcription. Moreover, Palmiter *et al.* described the brain as being lowest in the hierarchy of tissue-specific expression of a metallothionein-human growth hormone fusion gene (MT-hGH) introduced into transgenic mice⁶. Thus, it appears unlikely that increased transcription due to the MT-I promoter alone could account for the consistent elevation of HPRT fusion gene expression in the brain.

It is unknown what effect the HPRT cDNA sequence may have on the rate of fusion gene transcription. Charnay *et al.*⁷ and Wright *et al.*⁸ have demonstrated that the sequences responsible for differences in transcription of intact α - and β -globin genes are located 3' to the mRNA cap site of these genes, suggesting that *cis*-acting regulatory sequences lie within or beyond structural genes. Merrill *et al.* have shown that information sufficient to specify the regulation of the chicken thymidine kinase gene during cellular differentiation is entirely intragenic⁹. The expression of the MT-hGH and MT-hHPRT genes in the CNS of transgenic mice differs markedly: there is preferential expression of the MT-hGH gene in liver and low expression in the CNS. As MT-hHPRT behaves in a different manner, we conclude that the sequences within the hHPRT cDNA influence the elevated CNS HPRT expression.

Alternatively, increased CNS expression in these mice may involve tissue-specific differences in message stability. Leys *et al.* have demonstrated that expression of the gene encoding dihydrofolate reductase is controlled at the level of nuclear transcript stabilization¹⁰. Furthermore, Jefferson *et al.* have shown that the maintenance of high levels of liver-specific

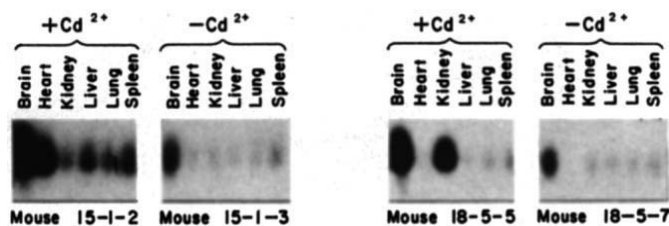


Fig. 4 Heavy-metal induction of HPRT fusion gene. Total RNA (20 µg) isolated from tissues of Cd²⁺-induced mice (15-1-2 and 18-5-5) and uninduced mice (15-1-3 and 18-5-7) was fractionated by electrophoresis through 1.1% agarose-formamide gels, transferred to nitrocellulose and hybridized with ³²P-labelled human HPRT cDNA. As this probe cross-hybridizes with endogenous mouse HPRT mRNA, Cd²⁺-responsiveness is demonstrated by signal intensification with RNA from treated animals.

mRNA molecules in cultured hepatocytes is a result of increased stabilization of these mRNAs¹¹. Presumably, information within these RNA molecules influences their selection for specific stabilization.

The present study suggests that sequences within the hHPRT cDNA determine increased HPRT specific activity in brain tissues. Whether the elevated expression of HPRT in CNS of transgenic mice or of normal mice and humans is the consequence of elevated transcription or message stability cannot be resolved from these studies; however, high levels of HPRT mRNA and enzyme activity in the CNS are observed in these animals in the absence of natural promoter or intron sequences.

We thank R. Palmiter for valuable advice and for the MT-hGH expression vector. We also thank Robert Hammer and Myrna Trumbauer for their contributions. This work was supported by the Howard Hughes Medical Institute, the Philip Michael Berolzheimer Medical Scientist Fellowship (J.T.S.), by NIH grant AM 31428 (to C.T.C.) and by NIH grant HD 17321 (to R.L.B.).

Received 15 January; accepted 29 July 1985.

- Seegmiller, J. E., Rosenbloom, F. N. & Kelley, W. N. *Science* **155**, 1682-1684 (1967).
- Kelley, W. N. *et al. Ann. Intern. Med.* **70**, 155-206 (1969).
- Howard, W. J., Kersen, L. A. & Appel, S. H. *J. Neurochem.* **17**, 121-123 (1970).
- Palmiter, R. D., Chen, H. Y. & Brinster, R. L. *Cell* **29**, 701-710 (1982).
- Durnam, D. M. & Palmiter, R. D. *J. Biol. Chem.* **256**, 5712-5716 (1981).
- Palmiter, R. D., Norstedt, G., Gelinas, R. E., Hammer, R. E. & Brinster, R. L. *Science* **222**, 809-814 (1983).
- Charnay, P. *et al. Cell* **38**, 251-263 (1984).
- Wright, S., Rosenthal, A., Flavell, R. & Grosfeld, F. *Cell* **38**, 265-273 (1984).
- Merrill, G. F., Hauschka, S. D. & McKnight, S. L. *Molec. cell. Biol.* **4**, 1777-1784 (1984).
- Leys, E. J., Crouse, G. F. & Kellems, R. E. *J. Cell Biol.* **99**, 180-187 (1984).
- Jefferson, D. M., Clayton, D. F., Darnell, J. E. & Reid, L. M. *Molec. cell. Biol.* **4**, 1929-1934 (1984).
- Brinster, R. L. *et al. Cell* **27**, 223-231 (1981).
- Southern, E. M. *J. molec. Biol.* **98**, 503-517 (1975).
- Chasin, L. & Urlaub, G. *Somat. Cell Genet.* **2**, 453-467 (1976).
- Johnson, G. G., Eisenberg, L. R. & Migeon, B. R. *Science* **203**, 174-176 (1979).
- Olsen, A. & Milman, G. *J. Biol. Chem.* **249**, 4030-4037 (1974).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Kafatos, F. C., Jones, W. C. & Efstratiadis, A. *Nucleic Acids Res.* **7**, 1541-1552 (1979).

Light-suppressible, cyclic GMP-sensitive conductance in the plasma membrane of a truncated rod outer segment

King-Wai Yau & Kei Nakatani

Department of Physiology and Biophysics, The University of Texas Medical Branch, Galveston, Texas 77550, USA

Recent experiments by Fesenko *et al.*¹ and ourselves² have shown that excised membrane patches from retinal rod outer segments contain a cyclic GMP-sensitive conductance which has electrical properties similar to those of the light-sensitive conductance. This finding supports the notion that cGMP mediates phototransduction (see ref. 3) by directly modulating the light-sensitive conductance. However, some uncertainty remained about whether the patch experiments had discriminated completely between plasma and

intracellular disk membranes⁴; thus the cGMP response in an excised membrane could have resulted from contaminating disk membrane fragments, which are known to contain a cGMP-regulated conductance⁵⁻⁸. Furthermore, the patch conductance has not yet been shown to be light-suppressible, an ultimate criterion for identity with the light-sensitive conductance. We now report experiments on a truncated rod outer segment preparation which resolved these issues. The results demonstrated that the cGMP-sensitive conductance was present in the plasma membrane of the outer segment, and that in the presence of GTP the conductance could be suppressed by a light flash. With added ATP, the effectiveness of the light flash was reduced and the suppression was more transient. The effects of both GTP and ATP were consistent with the known biochemistry (see refs 9-11). From the maximum current inducible by cGMP, we estimate that ~1% of the light-sensitive conductance is normally open in the dark; this would give an effective free cGMP concentration of a few micromolar in the intact outer segment in the dark.

Membrane current was recorded in darkness from a toad rod outer segment, either isolated or still attached to an inner segment, by drawing it partially into a suction pipette containing Ringer's solution¹². The inner segment and the basal portion of the outer segment outside the pipette were then broken off with a microprobe; this left an open-ended outer segment 'envelope' whose interior could be dialysed with a pseudointracellular solution in the bath. The potential difference between the pipette interior and the bath was held at 0 mV.

The upper trace in Fig. 1a shows a recording from an outer segment still attached to the inner segment. An inward current was present in the dark and was completely suppressed by a bright flash; this light response with its normal polarity indicated that the plasma membrane was still intact. After 'truncation', that is, removal of the inner segment and the basal portion of the outer segment, a large inward current developed on application of cGMP (Fig. 1a, lower trace). As in excised-patch experiments^{1,2}, this cGMP-induced current was little affected by large changes in Ca^{2+} concentration (Fig. 1b). The delays in the turning on and off of the current during changes in cGMP probably reflect the time required for cGMP equilibration in the outer segment, as they were significantly reduced for shorter truncated outer segments (see Fig. 3) and became unnoticeable in excised-patch experiments¹³. Occasionally (<10% of trials) a truncated outer segment showed little or no response to cGMP, perhaps because of resealing at the truncated end. When this happened, a sizeable cGMP-induced response could often be produced by pushing the outer segment out of the pipette for several micrometres and repeating the truncation.

The magnitude of the current induced by a saturating cGMP concentration depended on the length of the outer segment (L) within the recording pipette. In 9 of 21 experiments, such as that shown in Fig. 1c, the dependence was reasonably linear, indicating a uniform current density along most of the outer segment. This linear dependence was rather similar to that observed previously for the light-sensitive current¹². In the remaining experiments, however, the dependence was linear only with short L ($\leq 25 \mu\text{m}$); with longer L the current density declined progressively, perhaps because cGMP was unable to fully reach and open the conductance at the tip of the outer segment. With short L the longitudinal current density ranged from 4 to 20 $\text{pA } \mu\text{m}^{-1}$ at 0 mV. Assuming that healthier preparations showed more cGMP-responsiveness, we consider the top end of this range, 20 $\text{pA } \mu\text{m}^{-1}$, to be a more realistic estimate of the maximally induced current density.

Figure 2 shows the dependence of the current on cGMP concentration. To study this relationship, the outer segment was drawn into the pipette for 25 μm or less to minimize any incompletely dialysed space within the tip of the outer segment. The dose-response relation (Fig. 2b) was very similar to that found previously with excised patches of rod outer segment membrane^{1,2}. The current saturated at a cGMP concentration of ~1 mM. The sigmoidicity indicated positive cooperativity of cGMP molecules in turning on the conductance, with a Hill

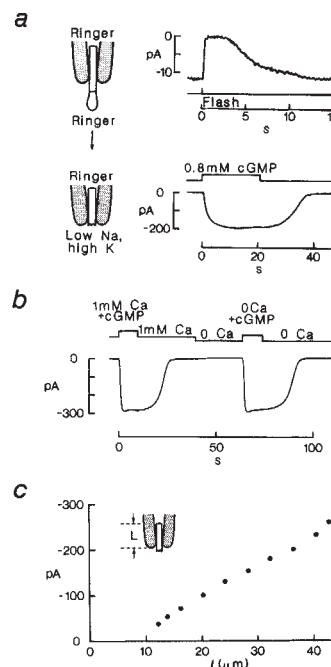


Fig. 1 Cyclic GMP-sensitive current recorded from a dialysed, open-ended rod outer segment of toad, *Bufo marinus*. **a**, Demonstration of the method. The length of outer segment within the pipette was 45 μm in both upper and lower panels. The flash delivered ~22 photons μm^{-2} (500 nm); diffuse unpolarized illumination. **b**, Changing the bath Ca concentration had little effect on the cGMP-sensitive current. The 0- Ca solution was identical in composition to the pseudointracellular solution described below; the 1 mM Ca solution contained 1 mM CaCl_2 and no TMA-EGTA, and was otherwise identical to the pseudointracellular solution. Length of outer segment within the pipette was 42 μm . cGMP concentration 0.84 mM. **c**, Dependence of recorded cGMP-sensitive current on the length of outer segment within the pipette. cGMP concentration 1.0 mM. Little change in seal resistance was observed throughout the experiment. The slope of the relation gave a current density of ~7.0 $\text{pA } \mu\text{m}^{-1}$. The data points did not cross the origin, probably because the point of maximal constriction, and hence the initial current-collecting point, of the pipette was some distance behind the tip¹². **Methods.** The outer segments, either isolated or still attached to inner segments, were obtained by finely chopping up a piece of dark-adapted retina in infrared light and then gently dispersing the chopped pieces. The suction pipette, which contained Ringer's solution, was connected to a current/voltage converter for recording membrane current¹². Seal resistance at the pipette tip was generally ~10 M Ω . Room temperature, 20-25 °C; 100-Hz low-pass (8-pole Butterworth) filtering. Normal Ringer's solution contained (in mM): 110 NaCl, 2.5 KCl, 1.6 MgCl_2 , 1.0 CaCl_2 , 5 TMA(tetramethylammonium)-HEPES, 5 dextrose, pH 7.6. Pseudointracellular solution contained (in mM): 12.5 NaCl, 100 K-gluconate, 1.6 MgCl_2 , 0.1 TMA-EGTA, 5 TMA-HEPES, pH 7.6. In control experiments with an intact, isolated segment, the addition of 1 mM cGMP produced a junction current²⁷ of ~3.5 pA; at lower cGMP concentrations the junction currents decreased correspondingly. For simplicity we have not corrected for these junction currents, but as the cGMP-sensitive current was generally much larger than the corresponding junction current this should make little difference except perhaps at 10 μM cGMP. Except for the experiment of Fig. 2, the time required for the bath solution to change was ~300 ms²⁸. The cGMP-containing solutions in **a-c** here and in Fig. 2 also contained 100 μM IBMX (isobutylmethylxanthine) to inhibit any phosphodiesterase activity. Different outer segments in **a-c**.

coefficient, n , of 2.5 and a half-saturating cGMP concentration, $K_{1/2}$, of 42 μM (Fig. 2c, solid circles). Based on averaged data from 12 experiments (Fig. 2c, open triangles), n was ~2.5 and $K_{1/2}$ ~45 μM . The close similarity between these results and those from the excised-patch experiments^{1,2} suggested that the same cGMP conductance was studied in both cases. In five

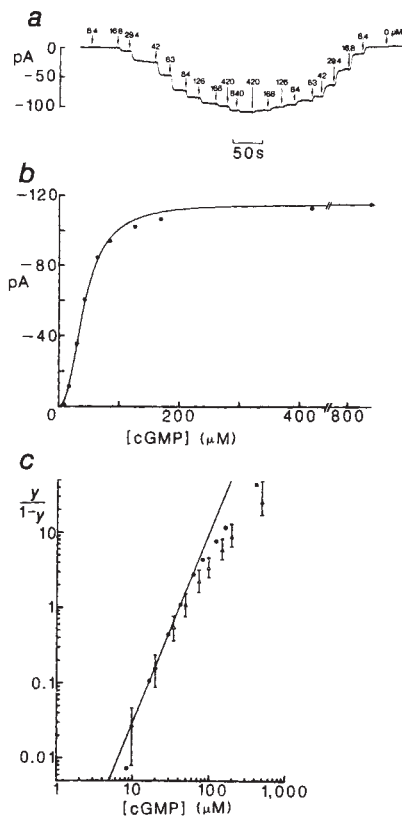


Fig. 2 Dependence of cGMP-sensitive current on cGMP concentration. *a*, Actual recordings. The arrows indicate the time points at which the cGMP concentration was changed. The time taken for a solution change in this experiment was ~2 s. Length of outer segment within the pipette was 19 μm . *b*, Linear plot of the averaged data from *a*. *c*, Hill plot of the data in *a* ($\bullet$), and of averaged results from 12 experiments (Δ). Error bars indicate standard deviations. y , Fraction of conductance open. In eight of the 12 experiments, the cGMP concentrations used were 10, 20, 35, 50, 75, 100, 150, 200, 500 and 1,000 μM . In the remaining four experiments, the cGMP concentrations were as shown in *a*. To average over all experiments, the data in the latter four experiments were first interpolated to 10, 20 μM , etc. before grand averaging was done. The smooth curve and the straight line fitted to the solid circles in *b* and *c* respectively were drawn according to the function $y(C) = C^n / [C^n + K_{1/2}^n]$ with $K_{1/2} = 42 \mu\text{M}$ and $n = 2.5$, where C is cGMP concentration.

experiments we have also made measurements both in the dark and in continuous bleaching light, and have observed no light-induced changes in n , $K_{1/2}$ or the saturated current (but see below). Thus, neither the characteristics of cGMP binding nor the intrinsic conduction properties of the light-sensitive conductance in the open state seemed to be influenced by light. In the Hill plot of Fig. 2c, there was systematic deviation of the data from a straight line at high concentrations of cGMP; this was also observed in excised-patch experiments (L. W. Haynes, K.N. and K.-W.Y., unpublished), and suggested that the binding of cGMP molecules to the receptor was probably not a single-step reaction (that is, $R + n \text{ cGMP} \rightleftharpoons R \cdot \text{cGMP}_n$) (see ref. 14).

The light-insensitivity of the cGMP-induced current was observed only in the absence of GTP (see, for example, Fig. 3b). With GTP in the bath, however, the current could be suppressed by light (Fig. 3a). In the presence of both GTP and ATP, the effectiveness of a light flash was reduced and the suppression also became more transient (Fig. 3c). Both the requirement for GTP in generating the light response and the quenching effect of ATP are consistent with present knowledge of the mechanisms controlling light-activated cGMP hydrolysis (see, for example, refs 9–11). The light responses in Fig. 3c resembled those in an intact cell¹², but they required brighter

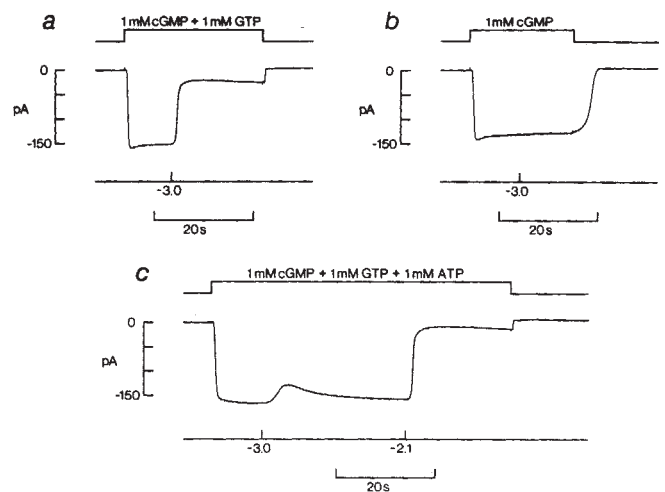


Fig. 3 Effects of GTP and ATP on the cGMP-sensitive current. Same outer segment for *a*–*c*; length of outer segment within the pipette was 25 μm . No IBMX was present in the cGMP-containing solutions. *a*, In the presence of GTP the cGMP-induced current became light-suppressible. *b*, The light-sensitivity disappeared on removal of GTP. *c*, Addition of GTP plus ATP restored the light response, but the sensitivity to light was significantly reduced. The number below each flash monitor indicates the nominal log attenuation of the light intensity. Diffuse unpolarized illumination. The '–3.0' flash delivered $\sim 56 \text{ photons } \mu\text{m}^{-2}$ (500 nm), which is close to an intensity that would normally just saturate the response in an intact toad rod (compare with Fig. 1a). The '–2.1' flash delivered $\sim 361 \text{ photons } \mu\text{m}^{-2}$ (500 nm).

light and had slower kinetics. Some quantitative differences compared with the responses of intact cells were perhaps inevitable because of the rather arbitrary nucleotide concentrations used in the experiment, and also because the dialysing solution most probably had disturbed some of the soluble and peripheral proteins involved in the light-controlled cGMP metabolism.

The results described here, together with the excised-patch experiments^{1,2}, provide conclusive evidence that the light-sensitive conductance in rods is modulated by cGMP. As the observed dependence of current on cGMP concentration is nonlinear, it might seem surprising that the intensity–response relation in an intact cell is linear at low light intensities¹². However, in normal conditions the visual excitation triggered by a photon is spatially confined to only a small region of the outer segment, within which the light-sensitive conductance is effectively shut down completely¹⁵. Consequently, as pointed out by Lamb *et al.*¹⁵ the intensity–response relation is determined simply by the Poisson distribution of photon arrival on the outer segment, which predicts a linear relation at low light intensities.

One parameter important for understanding the transduction mechanism is the free cGMP concentration present in the intact rod outer segment in the dark. Previous work has shown that the light-sensitive conductance of a rod is mostly closed in darkness^{22,29,30}, and that the dark current can increase 10–20-fold in conditions of low external or internal Ca concentration^{16,17} or high internal cGMP concentration^{17–19}. This suggests that about 5% (ref. 16) of the conductance is normally open in darkness. From the Hill plot of Fig. 2c, this would correspond to a free cGMP concentration of $\sim 13 \mu\text{M}$. This value, however, represents the upper estimate because the conductance in an intact cell probably cannot be made fully open owing to tight regulation of the cGMP metabolism^{20,31}.

The experiments described earlier indicated that the maximum current density inducible by cGMP was as high as $20 \text{ pA } \mu\text{m}^{-1}$ at 0 mV, equivalent to a total current of $\sim 1,000 \text{ pA}$ for an outer segment 50–60 μm long. From the current–voltage relation measured in intact cells as well as in excised patches, the current at a membrane potential of -40 mV is ~ 2.5 times that at 0 mV (refs 21, 22 and L. W. Haynes, K.N. and K.-W.Y., unpublished); this gives a maximum dark current of $\sim 2,500 \text{ pA}$ at rest, com-

pared with a normal current of ~25 pA. In other words, the conductance in the dark may be only ~1% open, giving a free cGMP concentration of ~6 μ M. From this or the above estimate, the free cGMP content in the dark would constitute only 5–10% of the total, if we take the total cGMP concentration within the rod outer segment to be ~60 μ M^{23–26} and the cytosolic space to be roughly half of the total volume. This may explain why a significant decline in cGMP content has not been detected in the intact retina in physiological lighting conditions²⁰. Based on the above estimates, even if all the free cGMP were hydrolysed in the light and the conductance completely shut down, the overall reduction of cGMP content in the outer segment would be only a few per cent.

We thank Drs Denis Baylor and Edward Pugh Jr for suggestions concerning the experiments and Drs Malcolm Brodwick and Lawrence Haynes for helpful comments on the manuscript. This work was supported by the National Eye Institute and the Retina Research Foundation, Houston, Texas.

Received 4 June; accepted 12 July 1985.

1. Fesenko, E. E., Kolesnikov, S. S. & Lyubarsky, A. L. *Nature* **313**, 310–313 (1985).
2. Nakatani, K. & Yau, K.-W. *Biophys. J.* **47**, 356a (1985).
3. Miller, W. H. (ed.) *Molecular Mechanisms of Photoreceptor Transduction* (Academic, New York, 1981).
4. Stieve, H. (ed.) *The Molecular Mechanism of Photoreception: Dahlem Konferenzen* (Springer, Berlin, in the press).
5. Caretta, A., Cavagioni, A. & Sorbi, R. T. *J. Physiol., Lond.* **295**, 171–178 (1979).
6. Caretta, A. & Cavagioni, A. *Eur. J. Biochem.* **132**, 1–8 (1983).
7. Koch, K. W. & Kaupp, U. B. *J. biol. Chem.* **260**, 6788–6800 (1985).
8. Puckett, K. L., Rawji, A. J. & Goldin, S. M. *Biophys. J.* **47**, 102a (1985).
9. Liebman, P. A. & Pugh, E. N. Jr, in *Molecular Mechanisms of Photoreceptor Transduction* (ed. Miller, W. H.) 157–170 (Academic, New York, 1981).
10. Kühn, H. in *Progress in Retinal Research* Vol. 3 (eds Osborne, N. N. & Chader, J.) 123–156 (Pergamon, New York, 1984).
11. Stryer, L. *Rev. Neurosci.* **9** (in the press).
12. Baylor, D. A., Lamb, T. D. & Yau, K.-W. *J. Physiol., Lond.* **288**, 589–611 (1979).
13. Haynes, L. W. & Yau, K.-W. *Nature* (submitted).
14. Dixon, M. & Webb, E. C. *Enzymes* (Academic, New York, 1979).
15. Lamb, T. D., McNaughton, P. A. & Yau, K.-W. *J. Physiol., Lond.* **319**, 463–496 (1981).
16. Yau, K.-W., McNaughton, P. A. & Hodgkin, A. L. *Nature* **292**, 502–505 (1981).
17. Matthews, H. R., Torre, V. & Lamb, T. D. *Nature* **313**, 582–585 (1985).
18. Cobbs, W. H. & Pugh, E. N. Jr *Nature* **313**, 585–587 (1985).
19. MacLeish, P. R., Schwartz, E. A. & Tachibana, M. *J. Physiol., Lond.* **348**, 645–664 (1984).
20. Goldberg, N. D., Ames, A. III, Gander, J. E. & Walseth, T. F. *J. biol. Chem.* **258**, 9213–9219 (1983).
21. Baylor, D. A. & Nunn, B. J. *J. Physiol., Lond.* (in the press).
22. Bodoia, R. D. & Detwiler, P. J. *J. Physiol., Lond.* (in the press).
23. Cohen, A. I., Hall, I. A. & Ferrendelli, J. A. *J. gen. Physiol.* **71**, 595–612 (1978).
24. Woodruff, M. L. & Bownds, M. D. *J. gen. Physiol.* **73**, 629–653 (1979).
25. Kilbride, P. & Ebrey, T. G. *J. gen. Physiol.* **74**, 415–426 (1979).
26. Woodruff, M. L. & Fain, G. L. *J. gen. Physiol.* **80**, 537–555 (1982).
27. Hodgkin, A. L., McNaughton, P. A., Nunn, B. J. & Yau, K.-W. *J. Physiol., Lond.* **350**, 649–680 (1984).
28. Yau, K.-W. & Nakatani, K. *Nature* **309**, 352–354 (1984).
29. Yoshikami, S. & Hagins, W. A. in *Biochemistry and Physiology of Visual Pigments* (ed. Langer, H.) 245–255 (Springer, New York, 1973).
30. Gray, P. & Attwell, D. *Proc. R. Soc. B* **233**, 379–388 (1985).
31. Yau, K.-W. & Nakatani, K. *Nature* **313**, 579–582 (1985).

Constitutive synthesis of interleukin-3 by leukaemia cell line WEHI-3B is due to retroviral insertion near the gene

S. Ymer, W. Q. J. Tucker, C. J. Sanderson†, A. J. Hapel*, H. D. Campbell & I. G. Young

Medical Molecular Biology Unit and *Department of Medicine and Clinical Science, John Curtin School of Medical Research, Australian National University, PO Box 334 Canberra, ACT, 2601 Australia

Interleukin-3 (multi-CSF) is a multilineage haematopoietic growth regulator that initiates the proliferation and differentiation of multipotential stem cells^{1–5}. Complementary DNA clones encoding interleukin-3 (IL-3) have recently been isolated^{6,7} and the structure of the IL-3 gene determined^{8,9}. IL-3 is produced by T lymphocytes or T lymphomas only after stimulation with antigens, mitogens or chemical activators such as phorbol esters. The myelomonocytic leukaemia line WEHI-3B^{10,11} also produces IL-3 but its production

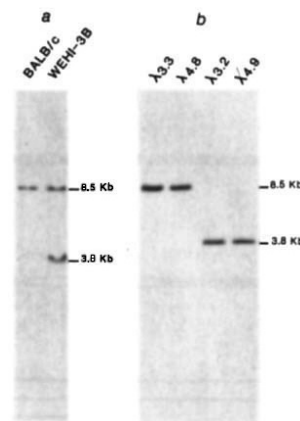


Fig. 1 Southern blots of genomic DNA and λ clones carrying the normal and rearranged IL-3 genes from WEHI-3B D⁺ cells. *a*, Genomic DNA from BALB/c mice and WEHI-3B cells; *b*, λ clones from a WEHI-3B genomic library. The DNA samples were digested with *EcoRI* restriction endonuclease, electrophoresed on a 1% agarose gel, transferred to nitrocellulose membrane and hybridized with a probe prepared from the *HindIII*–*NcoI* fragment of an IL-3 cDNA clone⁶ by random priming⁸.

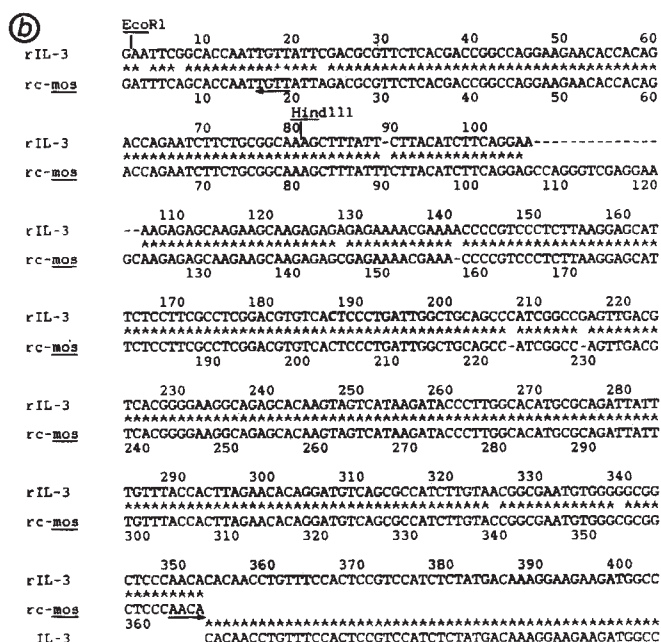
is constitutive^{12–14} and the WEHI-3B cells do not appear to produce significant levels of any of the other lymphokines normally secreted by T lymphocytes after stimulation. It has been proposed that the genetic change leading to the constitutive expression of IL-3 may have been a key event in the development of this leukaemia^{2,3,15,16}. We report here that the constitutive synthesis of IL-3 by the WEHI-3B cell line is due to the insertion of an endogenous retrovirus-like element close to the 5' end of the gene. The insertion, an intracisternal A particle (IAP) genome, is positioned with its 5' long terminal repeat (LTR) close to the promoter region of the IL-3 gene, resulting in constitutive synthesis of IL-3.

To investigate the reason for constitutive expression of the IL-3 gene in the leukaemia line WEHI-3B, genomic blots were carried out using a murine IL-3 cDNA probe^{6,8}. Since WEHI-3B cells were originally derived from BALB/c mice, genomic blots using BALB/c DNA were done in parallel. In BALB/c mice, blots of genomic DNA digested with *EcoRI* restriction endonuclease showed a single band of about 8.5 kilobases (kb) using an IL-3 cDNA probe (Fig. 1*a*), as expected^{8,9}. In the case of WEHI-3B, analogous blots showed the presence of this band plus a smaller band of similar intensity of about 3.8 kb. This suggested the presence of a normal IL-3 gene together with a rearranged IL-3 gene in WEHI-3B cells. The results of further blots carried out using genomic DNA digested with *BamHI* or *HindIII* restriction endonucleases suggested that the rearranged IL-3 gene differed from the normal gene only at its 5' end.

A genomic library was prepared from WEHI-3B DNA using the λ EMBL3A vector¹⁷, and screened¹⁸ for recombinant phage hybridizing to the IL-3 cDNA probe. These clones were found to belong to two different families corresponding to the normal and rearranged IL-3 genes inferred from the genomic blots, and possessed 8.5-kb and 3.8-kb *EcoRI* fragments respectively, which hybridized with the IL-3 cDNA probe (Fig. 1*b*). It was established that the restriction sites in the 3'-flanking region of the rearranged IL-3 gene were the same as those of the normal IL-3 gene from WEHI-3B cells. Digestion with *HpaII* and *HinfI* restriction endonucleases and hybridization with the IL-3 cDNA probe verified that both the normal and rearranged IL-3 genes from WEHI-3B possessed the fragments predicted from the IL-3 gene sequence⁸. Thus the rearrangement appeared to be localized to the 5'-flanking region of the IL-3 gene.

Digestion with *EcoRI*/*BamHI* and *EcoRI*/*HindIII* restriction enzymes and hybridization with both the IL-3 cDNA probe and a probe derived from the 5'-flanking region of the IL-3 gene from BALB/c mice indicated that in the rearranged gene, the normal 5'-flanking region of the IL-3 gene had been moved several kilobases upstream (Fig. 2). This provided clear evidence

† Permanent address: National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7 1AA, UK.



Methods. DNA sequencing was carried out by the chain termination method^{41,42} after subcloning appropriate fragments into M13mp8¹⁹. Sequencing was initiated using either a universal primer⁴³ or one of the three internal primers (P1, P2 or P3) 17 nucleotides long, which were synthesized by the phosphite triester method^{44,45}. The computer programs described by Staden⁴⁶ were used to analyse and assemble the sequence data.

The λ clones 3.2 and 3.3 carrying the rearranged and normal IL-3 genes respectively were transfected into BALB/c NIH 3T3 cells using the calcium phosphate precipitation method²⁴. Cultures were assayed periodically for IL-3 production using the IL-3-dependent cell line 32D-c123 (refs 25, 26). Cell transfected with λ clone 3.2 (carrying rIL-3) produced detectable IL-3 4 days post-transfection and the levels of IL-3 continued to increase up to day 8 (Fig. 5), reaching levels approximately 22% of those produced by WEHI-3B cells. In contrast, no IL-3 was produced by cells transfected with carrier DNA or transfected with λ clone 3.3 which carries the normal IL-3 gene. Clear verification of these results was obtained when the transient expression assays were repeated using monkey COS-1 cells. Six days after transfection, cells transfected with λ clone 3.2 gave

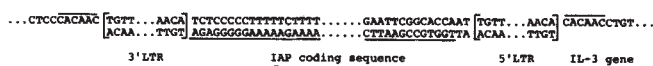


Fig. 4 Organization of the IAP genome in rIL-3. The 6-bp repeat of host sequence at the site of insertion of the IAP genome is shown together with the 4-bp inverted repeats which mark the extremities of each LTR. Apart from the 6-bp duplication at the site of insertion, no changes in the host sequence were found for 100 bp either side of the insertion site. The IAP genome is inserted in a head-to-head orientation, with the 5' LTR to the 5' end of the IL-3 gene. The IAP sequence adjacent to the 5' LTR (underlined) shows part of the 5' LTR tRNA primer binding site for (-) strand synthesis, and that adjacent to the 3' LTR shows the purine-rich (+) strand primer sequences²³.

IL-3 levels 63% of those found in WEHI-3B conditioned medium whereas those transfected with clone 3.3 produced no detectable IL-3.

The inability of the λ clone carrying the normal gene to be expressed efficiently in mouse 3T3 cells is in agreement with the silent nature of the endogenous IL-3 gene in these cells. These results are in contrast to those of Miyatake *et al.*⁹ who transfected the mouse IL-3 gene and pSV2neo into L cells and after selection for G418 resistance found that most of the transfected clones were producing low levels of IL-3. It is difficult to directly compare these experiments since the present work involves transient expression whereas the transfection experiments of Miyatake *et al.*⁹ involved selection for chromosomal integration. In addition, the expression of the integrated IL-3 genes may have been enhanced by adjacent copies of pSV2neo.

The IAP LTRs seem to have many features in common with other retroviral LTRs²⁷. In particular they appear to contain sequences for promotion, initiation, and termination of viral RNA transcription and for polyadenylation of the transcripts²³. They also have a conserved sequence GTGGTAAA similar to the simian virus 40 enhancer 'core' sequence GTGGAAA^{TTT} (refs. 21, 28). The insertion of the IAP genome 215 bases upstream of the IL-3 TATA box may have produced constitutive expression of IL-3 by one of several different mechanisms. The proximity of the 5' LTR to the IL-3 gene could provide a new promoter for the gene. The orientation of the 5' LTR is such that normal transcription from within the LTR would be directed away from the IL-3 gene, but there is formally the possibility of a cryptic promoter being present as has been demonstrated in the case of *rc-mos*²⁹. However, all three IL-3 cDNA clones from WEHI-3B which have been characterized⁶ have the same 5' end as the full-length cDNA clones derived by Yokota *et al.*⁷ from a T-cell line. Thus it seems likely that most transcripts in WEHI-3B are coming from the IL-3 promoter which is intact in the rearranged IL-3 gene. Another possibility is that the 5' LTR of the IAP genome is providing an enhancer sequence which promotes transcription of the IL-3 gene. Alternatively, the effect of the IAP genome insertion may have been to insertionally inactivate a *cis*-acting regulatory sequence or to move it upstream away from the IL-3 gene.

Intracisternal A particles are found budding from the endoplasmic reticulum in a variety of mouse tumours, in mouse embryos, but rarely in normal cells (see ref. 30). About 1,000 IAP genomes are present per haploid genome of *Mus musculus* and the IAP genomes, which are analogous to the proviral forms of retroviruses, vary in size from 4 to 7 kb. IAP transcripts have been shown to increase 50-fold over normal tissue in the case of a mouse colon tumour and to be significantly elevated in three different mouse leukaemias³¹. Insertional mutations by IAP genomes have been reported for κ light-chain genes²² where inactivation of gene function occurred and for the cellular oncogene *c-mos* in which case the gene was activated²⁹. In the present case insertion of a closely similar IAP genome to that present in *rc-mos* has resulted in constitutive expression of the IL-3 gene.

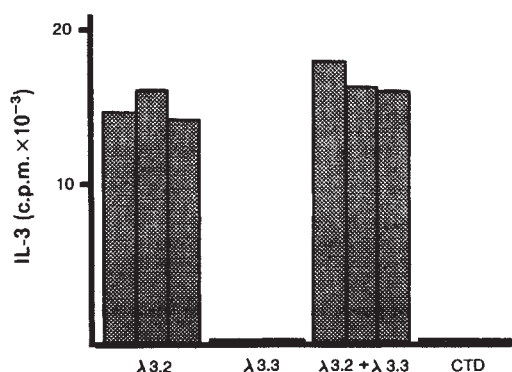


Fig. 5 Production of IL-3 by BALB/c 3T3 cells 8 days after transfection with λ clones from WEHI-3B (see Fig. 2) and carrier calf thymus DNA (CTD) as a control. IL-3 in the supernatants diluted 1:10 was assayed on the dependent cell line 32D-cl23 by measuring uptake of ³H-thymidine after 24 h^{25,26}. Cultures were transfected by the calcium phosphate technique²⁴ using 0.5 μ g of λ DNA and 0.5 μ g of CTD per culture. Data for three individual cultures (27 mm diameter) of 3T3 cells are shown in each case. Cultures transfected with mixed DNA from clones 3.2 and 3.3 produced similar levels of IL-3 to those transfected with clone 3.2 alone ($P > 0.05$), indicating that clone 3.3 had no inhibitory effect on IL-3 production. Counts obtained from cells transfected with clone 3.3 or CTD alone are not significantly different from controls in the IL-3 assay. Maximum counts obtained with WEHI-3B conditioned medium diluted 1:10 = 72,000 c.p.m.

There is growing interest in the role of autostimulatory mechanisms in oncogenesis^{32,33}. A variety of circumstantial evidence supports the possible autocrine role of transforming growth factors³⁴⁻³⁶ in certain cancers and several oncogenes appear to be related to growth factors or their receptors³⁷⁻³⁹. The growth and differentiation of cells of the haematopoietic system are regulated, at least *in vitro*, by specific growth factors¹. Any alterations in these normal regulatory mechanisms are of potential significance in the development of tumorigenicity. It is possible that WEHI-3B arose initially from an IL-3-dependent neutrophil-macrophage progenitor cell and that the activation of the IL-3 gene may have been an important transitional step in the production of this leukaemia. Further evolution of this progenitor may have resulted in bypassing of the initial autocrine loop. This may explain why WEHI-3B cells possess only low levels of IL-3 receptors (see ref. 40) although this could also be due to down-regulation.

We thank M. C. Fung and D. R. Cohen for helpful discussions, G. Mayo for synthesis of oligonucleotides, R. Johnson for preparing WEHI-3B DNA, A. George for providing λ packaging mix, K. Ridgway for preparing the EMBL3A vector and G. Quinn for programming assistance. This work was supported in part by the Children's Leukaemia and Cancer Foundation.

Received 31 May; accepted 26 July 1985.

- Nicola, N. A. & Vadas, M. *Immun. Today* **5**, 76-80 (1984).
- Iscove, N. N. & Roitsch, C. *Proc. IVth Int. Lymphokine Workshop* (Academic, New York, 1985).
- Garland, J. M. *Lymphokines* **9**, 153-200 (1984).
- Hapel, A. J. *et al. Blood* **65**, 1453-1459 (1985).
- Rennick, D. M. *et al. J. Immun.* **134**, 910-914 (1985).
- Fung, M. C. *et al. Nature* **307**, 233-237 (1984).
- Yokota, T. *et al. Proc. natn. Acad. Sci. U.S.A.* **81**, 1070-1074 (1984).
- Campbell, H. D., Ymer, S., Fung, M.-C. & Young, I. G. *Eur. J. Biochem.* **150**, 297-304 (1985).
- Miyatake, S., Yokota, T., Lee, F. & Arai, K.-I. *Proc. natn. Acad. Sci. U.S.A.* **82**, 316-320 (1985).
- Warner, N. L., Moore, M. A. S. & Metcalf, D. *J. natn. Cancer Inst.* **43**, 963-982 (1969).
- Metcalf, D. & Nicola, N. A. *Int. J. Cancer* **30**, 773-780 (1982).
- Metcalf, D., Moore, M. A. S. & Warner, N. L. *J. natn. Cancer Inst.* **43**, 983-1001 (1969).
- Ralph, P., Moore, M. A. S. & Nilsson, K. *J. exp. Med.* **143**, 1528-1533 (1976).
- Lee, J. C., Hapel, A. J. & Ihle, J. N. *J. Immun.* **128**, 2393-2398 (1982).
- Dexter, T. M. & Allen, T. D. *J. Path.* **141**, 415-433 (1983).
- Schrader, J. W. & Crapper, R. M. *Proc. natn. Acad. Sci. U.S.A.* **80**, 6892-6896 (1983).
- Frischauf, A.-M., Lehrach, H., Poustka, A. & Murray, N. *J. molec. Biol.* **170**, 827-842 (1983).
- Benton, W. D. & Davis, R. W. *Science* **196**, 180-182 (1977).
- Messing, J. & Vieira, J. *Gene* **19**, 269-276 (1982).
- Canaani, E. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 7118-7122 (1983).
- Christy, R. J., Brown, A. R., Gourlie, B. B. & Huang, R. C. C. *Nucleic Acids Res.* **13**, 289-302 (1985).
- Kuff, E. L. *et al. Proc. natn. Acad. Sci. U.S.A.* **80**, 1992-1996 (1983).

23. Ono, M. & Ohishi, H. *Nucleic Acids Res.* **11**, 7169-7179 (1983).
24. Graham, F. L. & Bacchetti, S. in *Techniques in the Life Sciences* Vol. B5 (Elsevier, Dublin, 1983).
25. Ihle, J. N. *et al. J. Immun.* **129**, 1377-1383 (1982).
26. Hapel, A. J., Warren, H. S. & Hume, D. A. *Blood* **64**, 786-790 (1984).
27. Varmus, H. E. *Science* **216**, 812-820 (1982).
28. Khoury, G. & Gruss, P. *Cell* **33**, 313-314 (1983).
29. Horowitz, M., Luria, S., Rechavi, G. & Givol, D. *EMBO J.* **3**, 2937-2941 (1984).
30. Kuff, E. L. & Fewell, J. W. *Molec. cell. Biol.* **5**, 474-483 (1985).
31. Augenlicht, L. M., Kobrin, D., Pavlovic, A. & Royston, M. E. *J. biol. Chem.* **259**, 1842-1847 (1984).
32. Heldin, C.-H. & Westermark, B. *Cell* **37**, 9-20 (1984).
33. Sporn, M. B. & Roberts, A. B. *Nature* **313**, 745-747 (1985).
34. Todaro, G. J. & De Larco, J. E. *Cancer Res.* **38**, 4147-5154 (1978).
35. Derynck, R., Roberts, A. B., Winkler, M. E., Chen, E. Y. & Goeddel, D. V. *Cell* **38**, 287-297 (1984).
36. Lee, D. C., Rose, T. M., Webb, N. R. & Todaro, G. J. *Nature* **313**, 489-491 (1985).
37. Waterfield, M. D. *et al. Nature* **304**, 35-39 (1983).
38. Doolittle, R. F. *et al. Science* **221**, 275-277 (1983).
39. Downward, J. *et al. Nature* **307**, 521-527 (1984).
40. Metcalf, D. *Science* **229**, 16-22 (1985).
41. Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
42. Sanger, F., Coulson, A. R., Hong, G. F., Hill, D. F. & Peterson, G. B. *J. molec. Biol.* **162**, 729-773 (1982).
43. Duckworth, M. L. *et al. Nucleic Acids Res.* **9**, 1681-1706 (1981).
44. Tanaka, T. & Letsinger, R. L. *Nucleic Acids Res.* **10**, 3249-3260 (1982).
45. McBride, L. J. & Caruthers, M. H. *Tetrahedron Lett.* **24**, 245-248 (1983).
46. Staden, R. *Nucleic Acids Res.* **8**, 3673-3694 (1980); **10**, 4731-4751 (1982); **12**, 499-503 (1984).

Expression of insulin-like growth factor-II transcripts in Wilms' tumour

Anthony E. Reeve*, Michael R. Eccles*,
Richard J. Wilkins*, Graeme I. Bell†
& Lynn J. Millow*

* Molecular Carcinogenesis Laboratory, Department of Biochemistry, University of Otago, PO Box 56, Dunedin, New Zealand

† Chiron Corporation, 4560 Horton Street, Emeryville, California 94608, USA

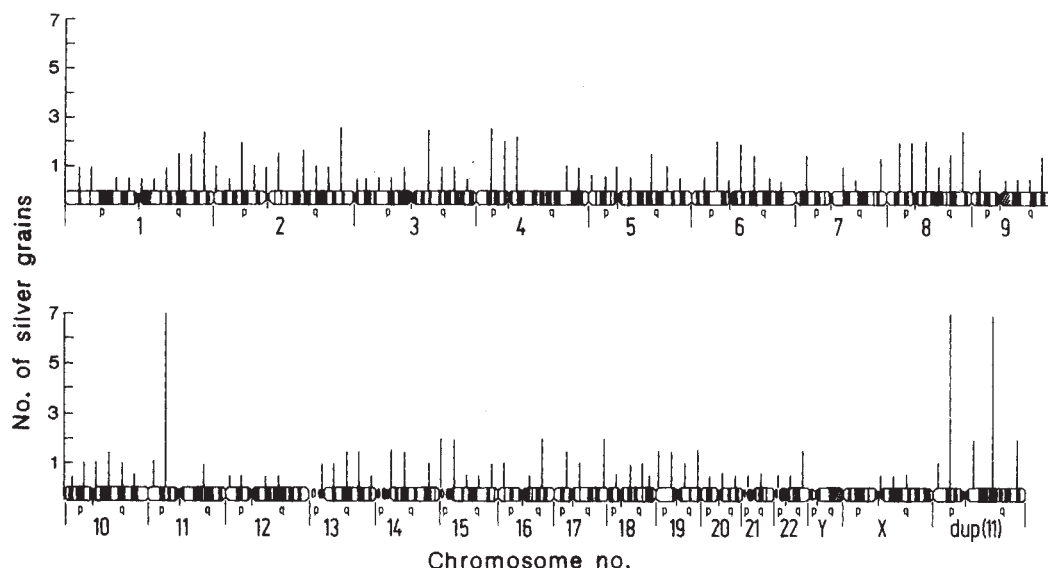
Wilms' tumour probably arises from embryonal kidney cells and occurs in both hereditary and sporadic forms¹. Knudson and Strong have suggested that both forms of the disease are initiated by two mutational events². In the case of the inherited form, cytogenetic evidence indicates that a germline deletion of chromosome band 11p13 may correspond to one of the two mutations³⁻⁵. DNA mapping evidence is consistent with the notion that the tumour susceptibility gene (*Wg*) on chromosome 11 is actually recessive⁶⁻⁹. Comings has proposed that the dominantly inherited tumours may arise by the inactivation or loss of a diploid pair of regulatory genes which normally suppress the expression of a structural transforming gene (*Tg*)¹⁰. It has recently been suggested that the *N-myc* oncogene may serve as a transforming gene in retinoblastoma¹¹, although no such gene has yet been identified in Wilms'

tumour. We now report that in four cases of Wilms' tumour, insulin-like growth factor-II (IGF-II) transcripts are highly elevated compared with the adjacent normal kidney. In addition, we have mapped the gene for IGF-II to chromosome band 11p14.1, which is in the immediate vicinity of *Wg*. These findings suggest that IGF-II may be involved in the aetiology of Wilms' tumour.

Initially we designed experiments to test whether the IGF-II locus is in the same 11p13 region as *Wg*. To do this, we used a lymphoblastoid cell line derived from a patient described as having a triplication of 11p material arising from the insertion of an 11p segment (11p11.2-11p14.1) into the long arm of one chromosome 11 at band 11q22.2 (refs 12, 13). This chromosome region encompasses *Wg*, and a half-sister of the above patient who is hemizygous for this segment has been described as having aniridia, multiple anomalies and mental retardation, which form part of the aniridia-Wilms' tumour association¹⁴.

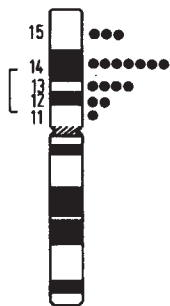
The chromosomal location of IGF-II was determined by the *in situ* hybridization of a full-length complementary DNA plasmid (phigf2)¹⁵ to metaphase chromosomes (cell line GM5913 from the Mutant Cell Repository, Camden, New Jersey) containing the 11p triplication. The phigf2 probe hybridized significantly ($P < 0.001$) to the short arms of the normal and duplicated chromosomes 11 (Fig. 1), thus confirming the recent assignments of IGF-II to the chromosome 11 short arm^{16,17}. In addition, a third hybridization site of equal intensity was detected on the long arm of the dup(11) chromosome but not on the long arm of the normal chromosome 11. As the position of the hybridization site on the dup(11) long arm corresponds to the site of insertion of the triplicated segment 11p11.2-11p14.1, we

Fig. 1 Distribution of IGF-II hybridization sites on chromosomes from the lymphoblastoid cell line GM5913. The numbers of silver grains observed over all the chromosomes (except chromosome 11 and dup(11)) were halved so that they could be compared directly with the haploid chromosome 11 and dup(11). The haploid karyotype has been divided into 145 equal segments and the total number of labelled sites was plotted for each segment. Hybridizations occurred significantly only to the triplicated region 11p11.2-11p14.1 ($P < 0.001$). Thirty metaphases were examined and half showed at least one grain on the triplicated region.



Methods. Mitotic chromosome preparations were obtained from methotrexate-synchronized cell cultures. The phigf2 plasmid was labelled to 5×10^8 d.p.m. per μg by nick-translation using ^{125}I -dCTP. The hybridization solution contained phigf2 ($300 \mu\text{g ml}^{-1}$), 10 mM sodium phosphate, 1 mM iododeoxycytidine, 10 mM potassium iodide, 300 mM sodium chloride, 30 mM sodium citrate, 50% formamide (v/v), 10% dextran sulphate (w/v) and $500 \mu\text{g ml}^{-1}$ heparin. *In situ* hybridization was done as previously described³². Slides were exposed to Kodak NTB-2 emulsion for 6-11 days, then developed and chromosomes were trypsin-Giemsa banded through the emulsion. The grain distribution was analysed using the Poisson distribution method³³.

Fig. 2 Distribution of silver grains on chromosome 11 after *in situ* hybridization with the IGF-II DNA probe. The phigf2 probe was hybridized to GM5913 chromosomes as described in Fig. 1 legend. The total number of grains hybridizing to the individual chromosome bands of both the short arms of the normal chromosomes 11 and dup(11) were scored.

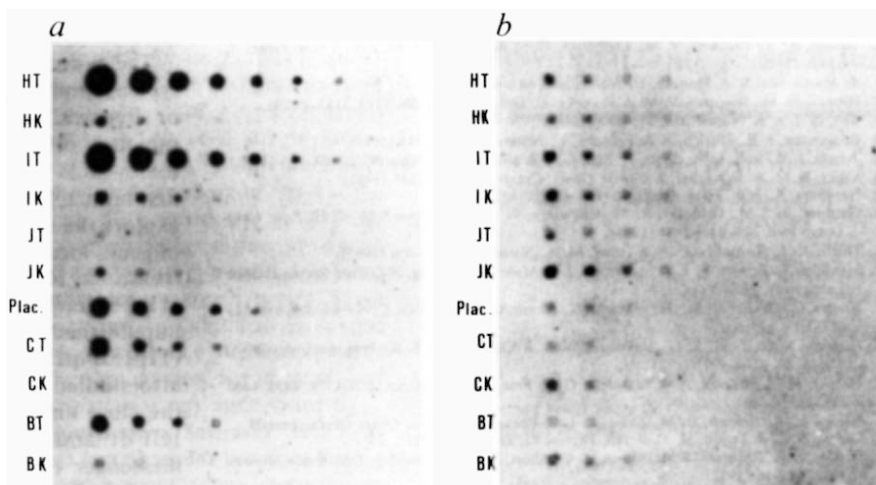


conclude that phigf2 hybridizes between 11p11.2 and 11p14.1.

Finer localization of the IGF-II gene was achieved by scoring the total number of grains over the individual bands on the short arm of the normal and duplicated chromosomes 11. The results were scored independently by two observers. The actual distribution of grains along the 11p arm was similar to a normal distribution, with a maximum near the boundary of bands 11p13 and 11p14 (data not shown). However, when grains were assigned to individual bands (Fig. 2), only band 11p14 hybridized significantly to the phigf2 probe. The apparent change in the distribution was due to the relative sizes of the bands. We conclude, therefore, that the IGF-II gene is located within the centromere-proximal region of sub-band 11p14.1, because it is within this band that the chromosome 11 break has been determined to occur during the 11p→11q insertion-translocation¹³.

The assignment of IGF-II to 11p14.1 is of interest because this segment is immediately adjacent to (and may even be within) the chromosome region which is deleted^{3-5,18,19} or involved in translocation²⁰ in Wilms' tumour. The close physical proximity of this growth-factor gene to *Wg* suggests that it may play a part in the aetiology of Wilms' tumour. To investigate this possibility, we have examined IGF-II RNA synthesis in sporadic Wilms' tumours from four patients, and have used as controls RNA preparations derived from adjacent histologically confirmed normal kidney. Total RNA samples from kidney and tumour were serially diluted, dot-blotted²¹ onto nitrocellulose and the nitrocellulose filter was hybridized with the phigf2 probe. The autoradiograph from this experiment showed clearly that IGF-II transcripts were 16–64-fold more abundant in the tumour samples than in the corresponding normal kidney (Fig. 3a, tumours H, I, B and C). A neuroblastoma of adrenal origin which had infiltrated the kidney was also analysed, but in this case IGF-II expression was not elevated (Fig. 3a, tumour J). A comparison of IGF-II expression levels between the tumour samples indicates that two were higher (H and I) while two were lower (B and C) than in a human term placenta (Fig. 3a). The significance of this variation is unclear, although the stage of the tumour or metabolic rate may contribute to this variation.

Fig. 3 RNA dot-blot analyses of Wilms' tumour and kidney RNA. Total RNA was isolated from tissue samples by the Gu SCN procedure³⁴. Tissue samples were obtained from tumours (T) and normal kidney adjacent to the tumour (K). Sample J was from a patient with neuroblastoma infiltrating the kidney, while all others were obtained from patients with Wilms' tumour (H, I, B and C). Tissue from a term placenta (Plac.) was also analysed. RNA samples were denatured in 0.44 M formaldehyde and 3 × SSC at 60 °C for 15 min, chilled on ice and twofold serial dilutions were applied to nitrocellulose sheets²¹. RNA was loaded from left to right using the following amounts: 5 µg, 2.5 µg, 1.2 µg, 0.6 µg, 0.3 µg, 0.15 µg, 0.07 µg and 0.03 µg, respectively. Blots were hybridized with either ³²P-labelled³⁵ phigf2 IGF-II insert (a) or a 2.9-kilobase *Sac*I fragment of the *Ha-ras* oncogene (b) which contains the coding regions³⁶. Blots were washed in 0.1 × SSC, 0.5% SDS at 68 °C.



These data nevertheless do suggest that in these Wilms' tumours, the level of IGF-II transcripts does not differ considerably from that observed in a normal expressing tissue. To determine whether placenta and tumour transcripts were similar, a Northern blot was prepared²² using IGF-II as the probe. The autoradiograph shown in Fig. 4 indicates that three major transcripts (6.0, 4.9 and 2.0 kilobases) were present in both placenta and tumour H. As anticipated from the dot-blot analyses (see above), IGF-II hybridization could not be detected to an identically prepared Northern blot of kidney RNA from this patient, thus eliminating any possible contribution to the hybridization signal of RNA sequences other than IGF-II. The RNA from tumours B, C and I was not completely intact and therefore a comparison was not made.

The *Ha-ras* gene is also known to be located on the short arm of chromosome 11 (refs 13, 23, 24), and the expression of this gene was analysed as a control on a duplicate of the dot-blot shown in Fig. 3a. In contrast to IGF-II, the abundance of *Ha-ras* transcripts varied only two- to fourfold between normal and tumorous tissue (Fig. 3b). It may be significant that *Ha-ras* expression is lower in tumours B and C than in normal kidney. Whether this reflects a lower tumour metabolic rate is unclear; however, as discussed above, these two tumours also had a lower level of IGF-II expression than tumours H and I. Further controls were done to examine the possibility that other insulin peptides were expressed in Wilms' tumour. Dot-blot analyses identical to that shown in Fig. 3 were done using insulin (chromosome 11 short arm²⁵) and IGF-I (chromosome 12 long arm^{16,17}) as probes. Insulin expression was below the level of detection in all samples, while IGF-I was expressed at lower levels in three tumours compared with the adjacent kidney (tumours I, B, C) and in one tumour (H) the level was double that in the corresponding kidney (data not shown). The transcription pattern exhibited by IGF-II in these Wilms' tumours was therefore quite distinct from that observed for the related IGF-I and insulin peptides.

While we cannot eliminate the possibility that IGF-II activation results from a subtle alteration within or surrounding the coding region, we can eliminate gross rearrangements and gene amplification, because Southern blot analyses of kidney and tumour DNA (*Hind*III digests) were identical (data not shown). When account is taken of our further observations that in three out of four of these tumours, chromosome 11 had undergone detectable changes involving a mitotic recombination (tumour H)²⁶, aneuploid loss (tumour I) or reduction to homozygosity (tumour C) (A. M. Raizis *et al.*, in preparation), the regulatory models of Knudson and Strong² and Comings¹⁰ appear most appropriate.

We therefore postulate that IGF-II expression occurs in Wilms' tumour as the result of both alleles of a regulator gene being inactivated or eliminated from the cell. Although we

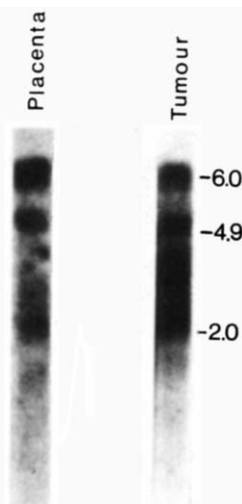


Fig. 4 Northern-blot analyses of Wilms' tumour and term placenta transcripts. Total RNA from placenta (20 µg) and tumour H (5 µg) was electrophoresed in a 1.5% agarose-formaldehyde gel²² and transferred to nitrocellulose filters. The blot was hybridized with ³²P-labelled phigf.2 IGF-II insert and washed as described in Fig. 3 legend. Transcript sizes were determined by comparison with a λ HindIII/EcoRI digest.

cannot exclude the possibility that IGF-II is simply an embryonal gene which is active at the developmental stage during which tumour initiation occurred, IGF-II does have many properties that are consistent with it being an embryonal transforming gene. For example, IGF-II is a mitogen that stimulates DNA synthesis in cultured embryo fibroblasts²⁷⁻²⁹. Furthermore, the rat counterpart of IGF-II (multiplication stimulating activity) is synthesized at high levels in the embryo but not during adult life, suggesting that it is important in the development of embryonal tissues^{30,31}. In addition, it is intriguing that IGF-II maps in the immediate vicinity of the presumptive *Wg* locus, and while the physical linking of regulatory and transforming genes is not an essential part of the carcinogenesis theory¹⁰, it does make some evolutionary sense for such genes to be linked if they are also required to interact during normal embryonal development.

We thank R. Gardner for advice during this work, G. Allen for help with chromosome preparation, D. Becroft, P. Crossen, P. Fitzgerald, J. Gillies and D. Holdaway for the tumour samples, and R. Weinberg for a gift of the pEJ6.6 plasmid. This work was supported by the Cancer Society of New Zealand, Otago Medical Research Foundation and the Otago Medical School Bequest Fund. M.R.E. is the Taranaki Cancer Research Fellow.

Received 28 January; accepted 29 July 1985.

- Gonzalez-Crussi, F. (ed.) & Franklin, W. A. in *Wilms' Tumour and Related Renal Neoplasms of Childhood*, 7 (CRC Press, Boca Raton, 1984).
- Knudson, A. G. Jr & Strong, L. C. *J. natn. Cancer Inst.* **48**, 313-342 (1972).
- Riccardi, V. M., Sujansky, E., Smith, A. C. & Francke, U. *Pediatrics* **61**, 604-610 (1978).
- Francke, U., Holmes, L. B., Atkins, L. & Riccardi, V. M. *Cytogenet. Cell Genet.* **24**, 185-192 (1979).
- Yunis, J. J. & Ramsay, N. K. C. *J. Pediatr.* **96**, 1027-1030 (1980).
- Koufos, A. *et al. Nature* **309**, 170-172 (1984).
- Orkin, S. H., Goldman, D. S. & Sallan, S. E. *Nature* **309**, 172-174 (1984).
- Reeve, A. E. *et al. Nature* **309**, 174-176 (1984).
- Fearon, E. R., Vogelstein, B. & Feinberg, A. P. *Nature* **309**, 176-178 (1984).
- Comings, D. E. *Proc. natn. Acad. Sci. U.S.A.* **70**, 3324-3328 (1973).
- Lee, W. H., Murphree, A. L. & Benedict, W. F. *Nature* **309**, 458-460 (1984).
- Strobel, R. J., Riccardi, V. M., Ledbetter, D. H. & Hittner, H. M. *Am. J. med. Genet.* **7**, 15-20 (1980).
- de Martinville, B. & Francke, U. *Nature* **305**, 641-643 (1983).
- Hittner, H. M., Riccardi, V. M. & Francke, U. *Ophthalmology* **86**, 1173-1183 (1979).
- Bell, G. I. *et al. Nature* **310**, 775-777 (1984).
- Brissenden, J. E., Ullrich, A. & Francke, U. *Nature* **310**, 781-784 (1984).
- Tricoli, J. V., Rall, L. B., Scott, J., Bell, G. I. & Shows, T. B. *Nature* **310**, 784-786 (1984).
- Slater, R. M. & de Kraker, J. *Cancer Genet. Cytogenet.* **5**, 237-245 (1982).
- Narahara, K. *et al. Hum. Genet.* **66**, 181-185 (1984).
- Gardner, R. J. M., Grindley, R. M., Chewings, W. E. & Holdaway, M. D. H. *Proc. Univ. Otago med. Sch.* **61**, 32-34 (1983).
- Thiele, C. J., Reynolds, C. P. & Israel, M. A. *Nature* **313**, 404-406 (1985).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. in *Molecular Cloning*, 202 (Cold Spring Harbor Laboratory, New York, 1982).
- Huerre, C., Despoisse, S., Gilgenkrantz, S., Lenoir, G. M. & Junien, C. *Nature* **305**, 638-641 (1983).
- Jhanwar, S. C., Neel, B. G., Hayward, W. S. & Chaganti, R. S. K. *Proc. natn. Acad. Sci. U.S.A.* **80**, 4794-4797 (1983).
- Harper, M. E., Ullrich, A. & Saunders, G. F. *Proc. natn. Acad. Sci. U.S.A.* **78**, 4458-4460 (1981).
- Raizis, A. M., Becroft, D. M., Shaw, R. L. & Reeve, A. E. *Hum. Genet.* (in the press).
- Dulak, N. C. & Temin, H. M. *J. cell. Physiol.* **81**, 153-160 (1973).
- Nissley, S. P., Rechler, M. M., Moses, A. C., Short, P. A. & Podskalny, J. M. *Endocrinology* **101**, 708-716 (1977).
- Moses, A. C., Cohen, K. L., Johnsonbaugh, R. & Nissley, S. P. *Clin. Endocr. Metab.* **46**, 937-946 (1978).

- Moses, A. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3649-3653 (1980).
- Adams, S. O., Nissley, S. P., Handwerger, S. & Rechler, M. M. *Nature* **302**, 150-153 (1983).
- Rabin, M. *et al. Cytogenet. Cell Genet.* **38**, 70-72 (1984).
- Bernheim, A., Berger, R. & Szabo, P. *Chromosoma* **89**, 163-167 (1984).
- Chirgin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5299 (1979).
- Rigby, P. W. J., Dieckmann, M., Rhodes, C. & Berg, P. *J. biol. Chem.* **113**, 234-251 (1977).
- Shih, C. & Weinberg, R. A. *Cell* **29**, 161-169 (1982).

Insulin-like growth factor-II gene expression in Wilms' tumour and embryonic tissues

J. Scott*, J. Cowell†, M. E. Robertson*, L. M. Priestley*, R. Wadey†, B. Hopkins‡, J. Pritchard†, G. I. Bell§, L. B. Rall§, C. F. Graham‡ & T. J. Knott*

* Molecular Medicine Research, MRC Clinical Research Centre, Harrow HA1 3UJ, Middlesex, UK

† Department of Molecular Genetics and Oncology, Institute of Child Health, University of London, Guilford Street, London WC1N 1EH, UK

‡ Department of Zoology, University of Oxford, South Parks Rd, Oxford OX1 3PS, UK

§ Chiron Corporation, 4560 Horton St, Emeryville, California 94608, USA

Wilms' tumour (nephroblastoma) is an embryonal neoplasm occurring in hereditary and spontaneous forms. Both types show rearrangements of the short arm of chromosome 11. The germ line of children with the rare inherited triad of aniridia, genitourinary abnormality and mental retardation carry a chromosome 11 that has a deletion in its short arm (band 11p13) and these children are at increased risk of developing Wilms' tumour^{1,2}. Neonates with the Beckwith-Wiedemann syndrome, in which there may be duplication of the 11p13-11p15 region, are similarly predisposed³. In the spontaneous form of the tumour a deletion of the 11p14 band in tumour cells, but not in normal cells, has been reported⁴, and the development of homozygosity for recessive mutations in the 11p region is implicated in the aetiology of Wilms' tumour⁵⁻⁸. In view of these chromosomal rearrangements and because Wilms' tumour is histologically indistinguishable from the early stages of kidney development⁹, we have now examined the expression of genes localized to 11p in Wilms' tumour and human embryonic tissue. In 12 sporadic tumours examined, the expression of the gene coding for insulin-like growth factor-II (IGF-II), localized to the 11p15 region, was markedly increased relative to adult tissues, but was comparable to the level of expression in several fetal tissues including kidney, liver, adrenals and striated muscle. This may reflect the stage of tumour differentiation, but could also contribute to the malignant process, as IGF-II is an embryonal mitogen¹⁰⁻¹³.

The Wilms' tumours studied were unilateral tumours of the typical spontaneous variety. The patients were without aniridia or Beckwith's syndrome. RNA was prepared from Wilms' tumours, other tumours of embryonal origin and adult tissue removed at surgery. Human fetal tissue was obtained from first-term therapeutic abortions and as full-term placenta. To explore the effect of the Wilms' tumour locus on the expression of genes localized to 11p, Northern and dot-blots were prepared from total RNA and hybridized with DNA probes for IGF-II (11p15) (refs 14, 15 and C. Morton, L.B.R., G.I.B. and T. Shows, unpublished data), insulin (11p15-p15.5)¹⁶⁻¹⁸, c-Ha-ras-1, (11p15.1-p15.5)¹⁸⁻²¹ and calcitonin (11p13-p15)²²⁻²⁴. We have also studied expression of the IGF-I (12q)¹⁵ gene because of the close amino-acid sequence relationship of this peptide to IGF-II and insulin, and have examined the production of messenger RNA for α-transforming growth factor (TGF), a developmentally regulated tumour mitogen²⁵ known to be produced in kidney (G.I.B., unpublished data).

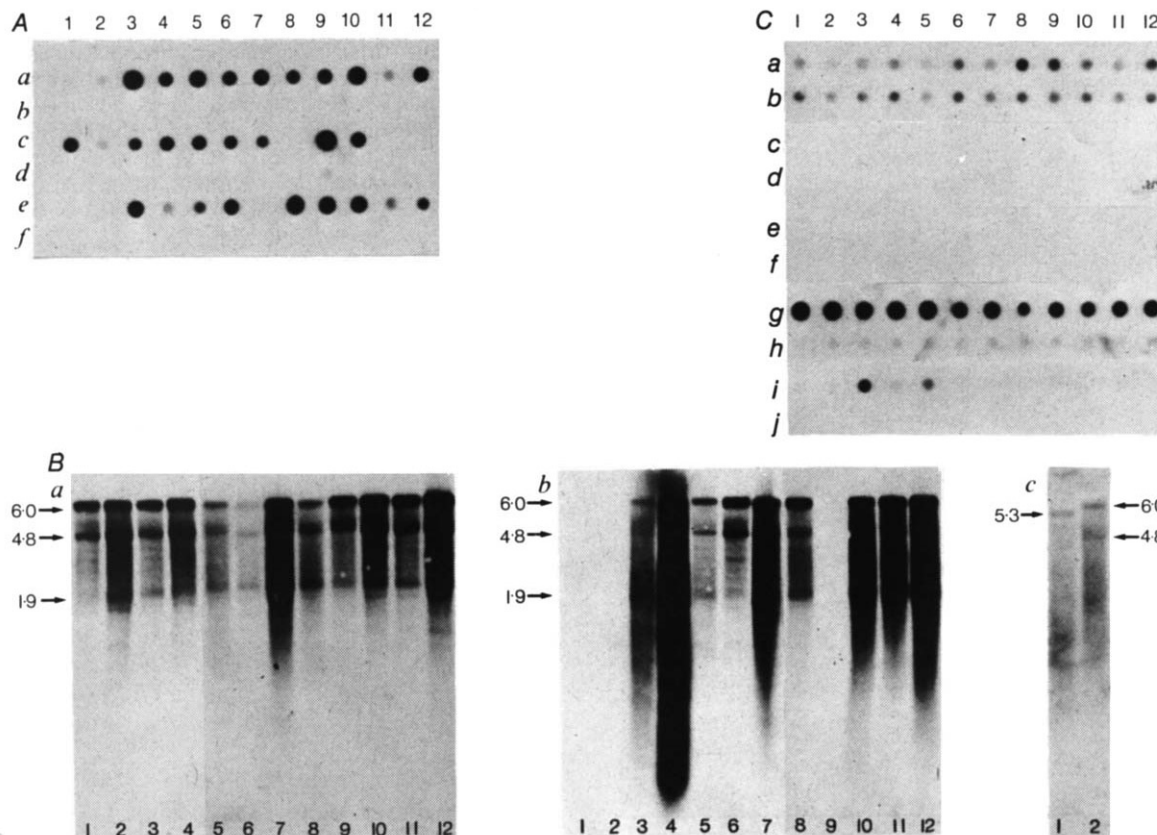


Fig. 1 **A**, Dot-blot analysis of total RNA (10 µg) isolated from tissues and probed with 32 P-labelled IGF-II cDNA. **a**, Lane 1, normal adult kidney; lanes 2–12, Wilms' tumour. **b**, Samples as **a**, but treated with NaOH as a control for hybridization to contaminating chromosomal DNA. **c**, Lane 1, Wilms' tumour; lane 2, adult liver; lanes 3–5, fetal kidney; lane 6, umbilical cord; lane 7, term placenta; lane 8, neuroblastoma; lanes 9, 10, rhabdomyosarcoma; lanes 11, 12, Ewing's sarcoma. **d**, Samples as in **c**, but treated with NaOH. **e**, Lane 2, L-cells; lane 3, fetal liver; lane 4, amnion; lane 5, villus trophoblast; lane 6, fetal tongue; lane 7, fetal brain; lanes 8–10, fetal adrenal; lanes 11, 12, fetal heart. **f**, Samples as **e**, but treated with NaOH. **B**, Northern blot analysis of total RNA (10 µg) hybridized to IGF-II cDNA. **a**, Lanes 1–12, Wilms' tumour. **b**, Lane 1, adult liver; lane 2, adult kidney; lanes 3, 4, rhabdomyosarcoma; lane 5, term placenta; lane 6, umbilical cord; lane 7, fetal tongue; lane 8, fetal striated muscle; lane 9, fetal brain; lane 10, fetal liver; lane 11, fetal kidney; lane 12, fetal adrenal. **c**, Lane 1, prolonged exposure of 10 µg adult human liver total RNA; lane 2, prolonged exposure of 40 µg of adult kidney RNA. Size markers were fragments of *Hind*III-digested λ DNA. Fragment sizes are in kb. **C**, Dot-blot analysis of total RNA (10 µg) hybridized with 32 P-labelled IGF-I cDNA (**a**, **b**). Insulin gene (**c**, **d**), calcitonin gene (**e**, **f**), c-Ha-ras-1 probe (**g**, **h**), and 925-base pair probe for entire coding region of α -TGF (**i**, **j**). The samples are: lane 1, **a**, **c**, **e**, **g**, **i**, normal adult kidney; lanes 2–12, Wilms' tumour; **b**, **d**, **f**, **h** and **j** are as **a**, **c**, **e**, **g**, **i**, but hydrolysed with NaOH. The NaOH-treated control for IGF-I shows that the signal is caused by contaminating DNA.

Methods. Surgical specimens were placed immediately in liquid nitrogen. Fetal samples were dealt with as described previously and frozen in liquid nitrogen²⁸. RNA was isolated, dot-blot and Northern blots prepared²⁹ using 10 µg of total RNA, hybridized with the appropriate DNA fragment 32 P-labelled by random priming³⁰ and washed.

In all 12 Wilms' tumours examined, IGF-II mRNA was increased 10–100-fold (Fig. 1A, Table 1) compared with adult liver and kidney, the principal sites of IGF-II gene expression in adult primates²⁶. Three major transcripts of 6.0, 4.8 and 1.9 kilobases (kb) were demonstrated in Wilms' tumour (Fig. 1B). In total RNA from normal adult kidney, similar-sized IGF-II transcripts were detectable. The only IGF-II transcript in adult liver was somewhat smaller at 5.3 kb. In contrast to adult tissues, adrenal, liver, kidney and striated muscle taken from 7–10-week human fetuses contained IGF-II mRNA concentrations that were comparable to those in Wilms' tumour, and the transcripts were of the same size as in the tumour. Slightly lower levels of these mRNAs were observed in first-term fetal heart, lung, intestine, umbilical cord, amnion, villus trophoblast, yolk sac and term placenta. The IGF-II mRNA sizes in these fetal tissues were also the same as in Wilms' tumour. The only fetal tissue examined that was negative for IGF-II gene expression was brain, in contrast to adult macaque²⁶ and human brain (J.S., M.E.R. and G. Taylor, unpublished data), where IGF-II mRNA is found in all regions. Examination of other embryonal tumours showed abundant IGF-II mRNA in rhabdomyosarcoma, a low level in Ewing's sarcoma of bone and embryonal carcinoma, but none in the neuroembryonal tumours neuroblastoma and retinoblastoma. Although the demonstration of IGF-II mRNA

in Wilms' tumour does not confirm that the protein itself is synthesized, the demonstration of similar mRNA concentrations and transcript sizes to those of fetal organs suggests that this is likely.

We found no expression of the genes for IGF-I, insulin and calcitonin in Wilms' tumour, fetal tissue (data not shown) or normal adult kidney (Fig. 1C). Comparable levels of c-Ha-ras-1 mRNA were present in normal kidney and Wilms' tumour. Low levels of α -TGF mRNA were found in four of the Wilms' tumours, but none was found in the fetal tissues examined (data not shown).

The IGF-II gene is located in the 11p15 region (C. Morton, L.B.R., G.I.B. and T. Shows, HGM8, 1985). The insulin and c-Ha-ras I genes are localized in the same band^{17–20}. This section of chromosome 11 is duplicated in some patients with Beckwith-Wiedemann syndrome who are characterized by the presence at birth of exomphalous, macroglossia and gigantism and who also have an increased risk of developing Wilms' tumour³. These features may result from overproduction of IGF-II. In view of the large physical distance between the Wilms' tumour locus at 11p13 and the IGF-II gene, it was not surprising to find no rearrangement or amplification of the gene in DNA isolated from these Wilms' tumours (data not shown). Two tumours studied had, however, developed homozygosity for chromosome

Table 1 Relative abundance of IGF-II mRNA in adult, fetal and tumour tissue

Tissue	Relative abundance		
	Adult	Fetal	Tumour
Liver	6	61	
Kidney	1	40	
Adrenal	—	60	
Small intestine	(0.08)	—	
Colon	(0.08)	—	
Unspecified region of intestine		6	
Spleen	(0.7)	—	
Lung	(0.2)	12	
Heart	(1.2)	18	
Skeletal muscle		38	
Tongue		49	
Brain	0.5	ND	
Term placenta		25	
Villus trophoblast		23	
Umbilical cord		35	
Amnion		10	
Yolk sac		19	
Wilms' tumour (nephroblastoma)			47 (6-98; 12)
Rhabdomyosarcoma			56, 102(2)
Neuroblastoma			ND(1)
Retinoblastoma			ND(1)
Ewing's sarcoma			1.3(2)
Teratocarcinoma			4(1-10; 4)
Hepatoblastoma (Hep G2 cells)			41
L-cells			ND

Relative levels of IGF-II were determined by averaging densitometer scans of several dot-blot matrices. The values are normalized to those of adult human kidney. Values in parentheses are for adult macaque²⁶. Data for tumours are shown respectively as the mean, range and number of tumours examined. The transcript sizes in Hep G2 cells were 6.0 and 4.8 kb. ND, none demonstrated.

11p markers (J.C., unpublished observations).

There are several possible explanations for the increased IGF-II gene expression in Wilms' tumour. (1) Deletion of a large region of 11p might lead to approximation of the Wilms' tumour locus to the IGF-II gene and consequent enhanced expression. While we cannot exclude this possibility, large deletions of a size necessary to produce this effect do not commonly occur in Wilms' tumour and we have not demonstrated altered expression of the insulin and c-H-ras genes which also reside at this locus¹⁷⁻²⁰. (2) The recessive mutation⁵⁻⁸ that leads to Wilms' tumour may cause re-activation of a set of embryonic genes which includes the gene coding for IGF-II, a growth factor widely produced during fetal development. Hence, the activation of this gene may reflect the stage of tumour differentiation. That the IGF-II gene is expressed at high levels in rhabdomyosarcoma and fetal muscle, but not in neuroembryonal tumours or fetal brain, supports this view. Whether increased IGF-II gene expression is associated with heterozygosity for the Wilms' tumour mutation and is responsible for the increased stature (R. S. Shannon, personal communication) and hemihyper trophy¹ that may be observed in Wilms' tumour patients, is unclear. (3) Since IGF-II is an embryonic mitogen¹⁰⁻¹³, the growth factor itself may initiate or contribute to the tumour process by autocrine or paracrine mechanisms²⁷. Finally, if the IGF-II protein itself is overproduced, it may serve as an important biochemical marker for the diagnosis of Wilms' tumour in those at risk or for monitoring the recurrence of the tumour in patients under treatment.

We thank Dr J. Heath, Dr L. M. Powell and C. Betsholtz for their contributions, Dr R. Craig for use of the calcitonin gene probe, Dr A. Hall for the c-Ha-ras-I probe, and Mrs A. Smith and Mrs T. Barrett for typing the manuscript. J.P. and J.C. receive support from the ICRF.

Note added in proof: A common pathogenetic mechanism involving the loss of heterozygosity for 11p markers has been

demonstrated in Wilms' tumour, rhabdomyosarcoma and hepatoblastoma³¹, the same embryonal neoplasms in which we demonstrate high IGF-II gene expression, and for which there is increased risk in Beckwith-Wiedemann syndrome. Embryonal tumours that remain heterozygous for 11p markers do not over-express the IGF-II gene. Our results further predict that embryonal tumours of adrenal origin, a fetal tissue that like kidney, muscle and liver also expresses the IGF-II gene at high levels, will arise by the same mechanism.

Received 14 May; accepted 24 July 1985.

- Miller, R. W., Fraumeni, J. F. & Manning, M. D. *New Engl. J. Med.* **270**, 922-927 (1964).
- Riccardi, V. M., Sujansky, E., Smith, A. C. & Francke, U. *Pediatrics* **61**, 604-610 (1978).
- Wazari, M., Patil, S. R., Hanson, J. W. & Bartley, J. A. *J. Pediatr.* **102**, 873-876 (1983).
- Kaneko, Y., Egues, M. C. & Rowley, J. D. *Cancer Res.* **41**, 4577-4578 (1981).
- Koufos, A. *et al. Nature* **309**, 170-172 (1984).
- Orkin, S. H., Goldman, D. S. & Sallan, S. E. *Nature* **309**, 172-174 (1984).
- Reeve, A. E. *et al. Nature* **309**, 174-176 (1984).
- Fearon, E., Vogelstein, B. & Feinberg, A. P. *Nature* **309**, 176-178 (1984).
- Mostofi, F. K. in *Kidney Disease: Present Status* (eds Churg, J., Spargo, B. & Mostofi, F. K.) 1-450 (Williams & Wilkins, Baltimore, 1979).
- Nissley, S. P., Rechler, M. M., Moses, A. C., Short, P. A. & Podskalny, J. M. *Endocrinology* **101**, 708-716 (1977).
- Moses, A. C., Cohen, K. L., Johnsonbaugh, R. & Nissley, S. P. *J. clin. Endocr. Metab.* **46**, 937-946 (1978).
- Moses, A. C. *et al. Proc. natn. Acad. Sci. U.S.A.* **77**, 3649-3653 (1980).
- Adams, S. O., Nissley, S. P., Handwerker, S. & Rechler, M. M. *Nature* **302**, 150-153 (1983).
- Bell, G. I. *et al. Nature* **310**, 775-777 (1984).
- Tricoli, J. V., Rall, L. B., Scott, J., Bell, G. I. & Shows, T. B. *Nature* **310**, 784-786 (1984).
- Bell, G. I., Karam, J. H. & Rutter, W. J. *Proc. natn. Acad. Sci. U.S.A.* **78**, 5759-5763 (1981).
- Harper, M. E., Ullrich, A. & Saunders, G. A. *Proc. natn. Acad. Sci. U.S.A.* **75**, 4458-4460 (1981).
- DeMartino, B. & Franke, U. *Nature* **305**, 641-643 (1983).
- Gerald, P. S. & Grzeschik, K. H. *Cytogenet. Cell Genet.* **37**, 103-126 (1984).
- Huerne, C., Despoisse, S., Gilgenkrantz, S., Lenoir, G. M. & Junien, C. *Nature* **305**, 638-641 (1983).
- Capon, D. J., Chen, E. Y., Levinson, A. D., Seeburg, P. H. & Goeddel, D. V. *Nature* **302**, 33-37 (1983).
- Craig, R. K., Hall, L., Edbrooke, M. R., Allison, J. & Macintyre, I. *Nature* **295**, 345-347 (1982).
- Hoppener, J. W. M. *et al. Hum. Genet.* **66**, 309-312 (1984).
- Prezepeiora, D. *et al. Biochem. biophys. Res. Commun.* **120**, 493-499 (1984).
- Todaro, G. J. *J. Cell Sci.* (In the press).
- Rall, L. B., Falloon, L., Styne, D. M. & Bell, G. I. *FEBS Lett.* (submitted).
- Sporn, M. B. & Todaro, G. J. *New Engl. J. Med.* **303**, 878-880 (1980).
- Thompson, S. *et al. J. Cell Sci.* **72**, 37-64 (1984).
- Rall, L. B. *et al. Nature* **313**, 228-231 (1985).
- Feinberg, A. P. & Vogelstein, B. *Analyt. Biochem.* **137**, 266-267 (1984).
- Koufos, A. *et al. Nature* **316**, 330-334 (1985).

A single genetic locus encoded by *Yersinia pseudotuberculosis* permits invasion of cultured animal cells by *Escherichia coli* K-12

Ralph R. Isberg & Stanley Falkow

Department of Medical Microbiology, Stanford University Medical Center, Stanford, California 94305, USA

For many species of pathogenic bacteria, invasion and survival within animal cells is central to establishing a successful host-parasite relationship¹⁻³. Localization within host cells protects the microorganism from host defences⁴, or permits it to cross epithelial barriers and subsequently become systemically distributed⁵. The precise mechanisms that permit entry of bacteria into host tissues are unclear⁶, therefore we have been studying the invasion of epithelial cells by *Yersinia pseudotuberculosis*^{7,8}. As a first step towards identifying the factors required for this process, we report here the identification of a single genetic locus from this organism that is sufficient to convert the innocuous *Escherichia coli* K-12 strain into an organism capable of invading cultured animal cells.

Sequences capable of conferring tissue invasion were identified in *E. coli* K-12 using the enrichment scheme illustrated in Fig. 1. This method should be capable of isolating genes involved in tissue invasion from any microorganism in which the necessary functions can be expressed and properly localized on the surface of *E. coli* K-12. In this procedure, a cosmid clone bank prepared from *Y. pseudotuberculosis* was introduced into an *E. coli* K-12 strain, as described elsewhere^{9,10}. The entire bank containing random sequences from *Y. pseudotuberculosis*

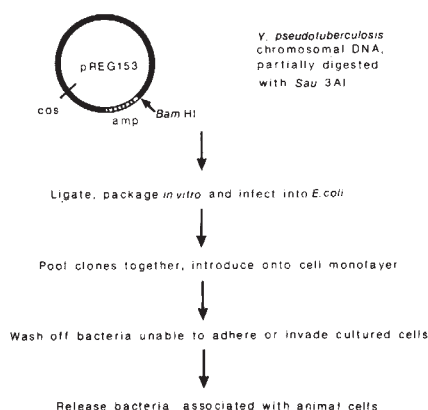


Fig. 1 The enrichment protocol for isolation of *E. coli* strains that are able to enter cultured animal cells. A cosmid bank was constructed^{9,10} by ligating a size-fractionated *Sau*3AI digest of chromosomal DNA isolated from *Y. pseudotuberculosis* strain YPIII(p⁻)¹⁷ into the *Bam*HI site of pREG153 (given by Dr R. Gill), a cosmid cloning vehicle that encodes resistance to ampicillin. Aliquots of the ligation mix were packaged *in vitro*¹⁸ into phage λ heads with an extract prepared from MMS A345 (given by Dr F. Stahl) and then infected into *E. coli* K-12 strain HB101 (ref. 19), followed by selection for ampicillin-resistant transductants. Pooled bacteria were grown overnight in culture, washed twice in sterile PBS, and resuspended to a density of 5×10^8 bacteria ml⁻¹. Two ml of the washed bacteria were then introduced onto a confluent monolayer containing 1×10^7 HEP-2 cells grown in RPMI 1640 medium (Irvine Biologicals), and incubated for 3 h at 36 °C in a 5% CO₂ atmosphere. Monolayers were then washed 10 times with sterile PBS to remove non-adherent bacteria. Bacterial clones remaining associated with the monolayer were released with 1% Triton X-100 in deionized water, and plated for individual colonies onto bacteriological medium.

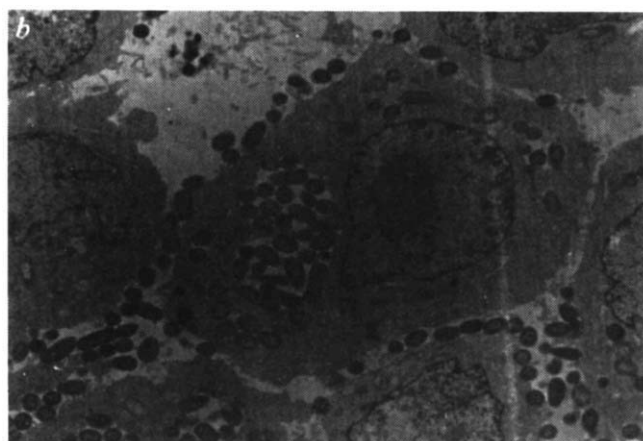
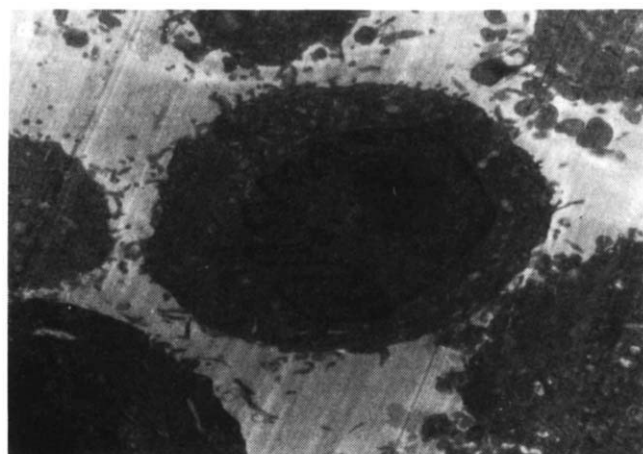


Fig. 2 Thin-section electron microscopy of cultured cells incubated with an *E. coli* strain harbouring the plasmid containing the *Y. pseudotuberculosis* invasion region. *a*, Thin section of HEP-2 cells infected with HB101 harbouring plasmid pRI203.14::Tn5, which is pRI203 (Table 1) containing a Tn5 insertion mutation in the *inv* locus. *b*, Identical to *a*, except HB101 harbours pRI203 with an intact *inv* locus.

Methods. Tissue culture wells (Falcon 3046, 6-well dishes) seeded with 8×10^5 HEP-2 cells in RPMI 1640 medium were incubated in the presence of 6×10^7 bacteria for 3 h at 36 °C. Monolayers were then washed 10 times with PBS and incubated for 10 min at 37 °C in the presence of PBS containing 0.1 mM EDTA. The monolayers were gently washed once more with PBS in the absence of EDTA, suspended in 1 ml of PBS, and pelleted at 600g for 10 min. The cell pellets were successively fixed with 2% glutaraldehyde and 2% osmium tetroxide in 0.1 M cacodylate buffer (pH 7.4), before staining with uranyl acetate as described¹⁴. Samples, dehydrated in ethanol, were embedded in Spurr's (Polysciences), thin-sectioned, stained successively with 1% uranyl acetate and lead acetate²⁰, and visualized with a Phillips 201c electron microscope.

teria appeared to be both bound to the outside of the cell and present within large endocytic vesicles. These micrographs, taken together with the gentamicin-protection assay described above, show that the invasive phenotype can be successfully mimicked by introducing DNA from a *Yersinia* species into *E. coli*.

Restriction enzyme analyses of plasmid DNA isolated from several of the bacterial clones that survived the enrichment procedure showed that these plasmids had identical or overlapping DNA sequences. To determine the physical location of the *inv* locus, one such plasmid, pINVA2, was analysed in detail. A series of Tn5 insertion mutations were isolated in this cosmid DNA, taking the precaution that each mutation analysed was the result of an independent event. As can be seen in Fig. 3, all the insertion mutations that eliminated the ability to enter cultured cells mapped in a contiguous 3.2-kilobase (kb) region.

was pooled, grown in broth culture, and used to infect a monolayer of cultured HEP-2 cells¹¹. After several hours of incubation, bacteria that were unable to associate stably with the cell monolayer were removed by washing. Bacteria remaining associated with the HEP-2 cells were isolated on an appropriate growth medium following treatment of the monolayer with Triton X-100, a gentle detergent (Fig. 1).

Twenty-two individual bacterial clones surviving the enrichment were then tested for internalization by the animal cells, by assaying for the ability of the bacteria to escape killing by an antibiotic that is unable to penetrate mammalian cells^{12,13}. Briefly, individual clones were introduced into microtitre wells containing HEP-2 monolayers (Table 1). After allowing for binding and entry of the bacteria into the cultured cells at 36 °C, the monolayers were washed free of unabsorbed bacteria, then incubated at the same temperature in fresh medium containing the antibiotic gentamicin. This aminoglycoside is unable to penetrate cultured mammalian cells, so if the *E. coli* cells containing the cloned *Y. pseudotuberculosis* DNA actually enter the HEP-2 cells, they should resist the bactericidal activity of the antibiotic. It was found that 12 of the 22 candidate strains survived the treatment with gentamicin. Data for several of the strains, shown in Table 1, indicate that a significant fraction of the bacteria escaped gentamicin killing. It was striking that the efficiency of escape from gentamicin treatment, which we presume represented bacterial invasion, was similar to that found for the *Y. pseudotuberculosis* strain used as the DNA donor. We will refer to the *Y. pseudotuberculosis* locus that encodes this property as *inv*.

To determine whether we had isolated bacterial derivatives that could invade cultured animal cells, ultrathin sections of monolayer cells exposed to several of the bacterial strains were analysed by electron microscopy (Fig. 2)¹⁴. As can be seen in the typical micrographs shown in Fig. 2, an *E. coli* K-12 strain is unable to bind or enter HEP-2 cells. In contrast, the same bacterial strain harbouring an intact *inv* locus showed a large number of bacteria associated with the animal cell. These bac-

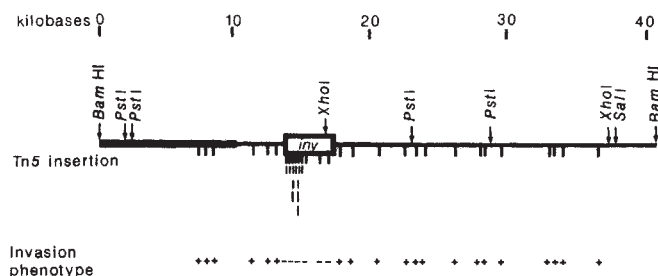


Fig. 3 Physical and genetic map of the *Y. pseudotuberculosis* locus that confers invasiveness on *E. coli* K-12. Shown is a map of pINVA2, a cosmid isolated following the enrichment described in Fig. 1 legend. Vertical lines below the map indicate the positions of Tn5 insertion mutations. —, Mutations eliminating invasiveness; +, mutations having no effect on invasiveness. Thickened line, DNA corresponding to pREG153 sequences. Open box, region in which all the invasion-defective mutations map. Only one of the *Bam*HI sites is shown in map.

Methods. Mutations induced by the kanamycin-resistance transposon Tn5 were isolated via transposition from λ 221*rex*:Tn5 c18570am23Pam80 onto a cosmid that confers invasiveness²¹. To select for insertions onto the cosmid, kanamycin-resistant colonies were pooled, and infected lytically with λ lac5 *imm*21*cts*. The resulting lysate was used to transduce HB101 to simultaneous kanamycin- and ampicillin-resistance. Such transductants contained insertions of Tn5 at random sites on the cosmid. Several hundred such transductants were assayed for invasiveness, as described in Table 1, and the map position of 20 mutants that eliminated invasiveness and 20 insertions that had no effect on the invasive phenotype were determined. To ensure that each insertion was independent, the physical location of no more than one mutation from each pool was analysed.

Therefore, it is apparent that relatively little genetic material from *Y. pseudotuberculosis* is sufficient to permit *E. coli* to enter epithelial cells. Tests with subclones from this region indicate that the synthesis of only one large protein from this region is required for the expression of the invasive phenotype (data not shown).

The *E. coli* strains described here display one of the cardinal pathogenic traits of several important enteric pathogens. Nevertheless, it should be emphasized that this phenotype alone is not sufficient to make the bacterium virulent, as is clear from studies on invasive *E. coli* strains by Sansonetti *et al.*⁶

There are several important ramifications from these data. First, by reconstructing the invasion phenotype of *Y. pseudotuberculosis*, we hope to exploit the techniques of bacterial genetics to dissect this important process. To this end, we have isolated and mapped 40 mutations that delimit the region required by *E. coli* for the entry process (Fig. 3). Second, it appears that this phenotype, at least in *Yersinia*, may be the result of a much simpler set of factors than had been assumed from the work on other invasive pathogens, especially *Shigella*¹⁵. It is striking that the only factors missing from *E. coli* that prevent it from invading cultured cells are encoded in a 3.2-kb region from *Y. pseudotuberculosis*. Third, this work may be a first step towards defining a previously undescribed class of proteins encoded by many invasive disease-producing microorganisms, as part of their strategy of pathogenesis. Determining how these proteins act may be useful in understanding an essential step in the infection process—and possible ways to prevent this step. Finally, the relative simplicity of this system immediately suggests that clones encoding factors for invasion might be exploited to introduce macromolecules, or other appropriate reagents, into animal cells for research and therapeutic purposes.

We thank Lynn Mercer for tutoring and advice on electron microscopic techniques and J. Michael Koomey, Carol Kumamoto and Pamela Small for critical evaluation of this work. R.R.I. was supported by the Jane Coffin Childs Fund for Medical Research. This research was supported, in part, by NSF grant PCM 83-06654 and contract DAMD 17-82-C-2002 from the US Army Medical Research Acquisitions Agency to S.F.

Table 1 Enrichment procedure yields *E. coli* strains that invade cell culture monolayers

Strain	% Invasion§
YPIII(p ⁻)*	9.0
HB101†	0.005
HB101(pINVA2)‡	8.3
HB101(pINVA7)	7.9
HB101(pINVG10)	8.7
HB101(pRI203)	9.2

Saturated cultures of bacteria, grown at 28 °C, were washed twice in phosphate-buffered saline (PBS) and resuspended to a concentration of 3×10^8 bacteria ml⁻¹. Aliquots of 50 μ l of each strain were added to monolayer cultures of HEP-2 cells seeded at a concentration of 2×10^5 animal cells per microtitre well (24-well, Falcon 3047 microtitre dishes) in RPMI 1640 medium. Bacteria were centrifuged onto the monolayer at 600g as described¹³, and the infected cultures were incubated at 36 °C for 3 h in a 5% CO₂ atmosphere, to allow binding and invasion of bacteria. Non-adherent bacteria were removed from the monolayer by washing three times with sterile PBS, and RPMI medium containing 40 μ g ml⁻¹ gentamicin (Sigma) was added to each microtitre well. The incubation was continued at 36 °C for 2 h in the presence of the antibiotic, before washing the monolayers in PBS twice more. Internalized bacteria were then released from the monolayers by the addition of 1% Triton X-100, and titred on L agar plates¹⁶.

* *Yersinia pseudotuberculosis* strain¹⁷.

† *E. coli* K-12 strain HB101.

‡ HB101 harbouring cosmids that are denoted in parenthesis.

§ Percentage of bacteria added to HEP-2 monolayers that resist treatment by gentamicin.

Received 16 April; accepted 9 July 1985.

- Formal, S. B., Hale, T. L. & Sansonetti, P. J. *Rev. infect. Dis.* **5**, S702-S707 (1983).
- Densen, P. & Mandell, G. L. *Rev. infect. Dis.* **2**, 817-838 (1980).
- Horwitz, M. A. *J. exp. Med.* **158**, 2108-2126 (1983).
- Armstrong, J. A. & D'arcy-Hart, P. J. *exp. Med.* **134**, 713-740 (1971).
- Une, T. *Microbiol. Immun.* **21**, 349 (1977).
- Sansonetti, P. J. *et al. Infect. Immunity* **39**, 1392-1402 (1983).
- Bovallius, A. & Nilsson, G. *Can. J. Microbiol.* **21**, 1997-2007 (1975).
- Bolin, I., Norlander, L. & Wolf-Watz, H. *Infect. Immunity* **37**, 506-516 (1982).
- Hohn, B. *Meth. Enzym.* **68**, 299-309 (1979).
- Koomey, J. M., Gill, R. E. & Falkow, S. *Proc. natn. Acad. Sci. U.S.A.* **79**, 7881-7885 (1982).
- Moore, A. E., Sabachewsky, L. & Toolan, H. W. *Cancer Res.* **15**, 598 (1957).
- Vesikari, T., Bromirska, J. & Maki, M. *Infect. Immunity* **36**, 834-836 (1982).
- Devenish, J. A. & Schiemann, D. A. *Infect. Immunity* **32**, 48-55 (1981).
- Horwitz, M. A. *J. exp. Med.* **158**, 1319-1331 (1983).
- Hale, T. L., Sansonetti, P. J., Schad, P. A., Austin, S. & Formal, S. B. *Infect. Immunity* **40**, 340-350 (1983).
- Miller, J. H. *Experiments in Molecular Genetics* (Cold Spring Harbor Laboratory, New York, 1972).
- Bolin, I. & Wolf-Watz, H. *Infect. Immunity* **43**, 72-78 (1984).
- Enquist, L. & Sternberg, N. *Meth. Enzym.* **68**, 281-298 (1979).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning: A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).
- Reynolds, E. S. *J. Cell Biol.* **17**, 208-213 (1963).
- deBruijn, F. J. & Lupski, J. R. *Gene* **27**, 131-149 (1984).

Growth control of activated, synchronized murine B cells by the C3d fragment of human complement

Fritz Melchers*, Anna Erdei*, Thomas Schulz† & Manfred P. Dierich†

* Basel Institute for Immunology, Grenzacherstrasse 487, Postfach, CH-4058, Basel, Switzerland

† Institut für Hygiene, University of Innsbruck, A-6010, Innsbruck, Austria

Three restriction points control the cell cycle of activated B lymphocytes¹. The first occurs directly after mitosis and is controlled by the occupancy of surface-bound immunoglobulin. The second is observed ~4 h after mitosis in the G₁ phase of the cycle, that is, before DNA replication, and is controlled by growth factors that are produced by macrophages² which we have previously classified as α -type factors^{3,4}. The third restriction point occurs in the G₂ phase, 2-4 h before mitosis, and is controlled by β -type

growth factors probably produced by helper T lymphocytes⁵. The third component of complement, C3, has long been implicated in the control of B-cell responses⁶⁻¹². C3 is secreted by monocytes and macrophages^{13,14}. We have found recently that crosslinked, but not soluble, human C3 stimulates activated, but not resting, murine B cells to thymidine uptake¹⁵. Here we investigate the role of C3b and C3d in the progression of the cell cycle of activated, synchronized murine B cells. We find that crosslinked C3d replaces the action of α -factors within the cell cycle of these cells and allows entry into S phase. In contrast, soluble C3d inhibits the action of α -factors. This implies that a C3d-specific receptor, probably the murine analogue to the human complement receptor CR2 (refs 16, 17), is a growth factor receptor on activated B cells that will give the cell a growth-positive signal when it is crosslinked, while occupancy by the soluble form of C3d will result in inhibition of the action of α -factors or of crosslinked C3b or C3d. A stretch of weak homology between the cDNA sequence of murine C3d and those of murine growth factors indicates that an insulin-like growth factor could be the active principle of C3d that controls the cell cycle of activated B cells.

First, the action of Sepharose-bound C3b and C3d was tested on activated, synchronized B cells from athymic *nu/nu* C57BL/6J mice. For activation, spleen cells were enriched for small resting cells by velocity sedimentation¹⁸, then activated in the presence of macrophage-derived α -factors^{3,4} by 2 days' treatment with lipopolysaccharide (LPS)¹⁹. The activated cells were again separated by velocity sedimentation and fractions of a given size were cultured in the presence of the various stimuli indicated in Figs 1 and 2. These activated cells were almost entirely B cells and devoid of α -factor-producing monocytes or macrophages^{1,20}. LPS-activated B cells of a given size underwent a further, synchronous division even in the absence of any stimulus, or in the presence of only one of the stimuli (Sepharose-bound immunoglobulin-specific antibodies, α -factors, β -factors, Sepharose-bound C3b, or Sepharose-bound C3d alone). This single division is probably caused by cell-bound LPS with which the cells have been activated from the resting state.

When α - and β -factors were added together, the activated B cells divided once more, while in the added presence of Sepharose-bound immunoglobulin-specific antibodies, they underwent several (at least three) synchronous divisions (Fig. 1a). Sepharose-bound C3b or Sepharose-bound C3d was then added to substitute for either α - or β -factors in their action on activated, synchronized B cells. Although neither replaced the action of β -factors, both could replace α -factors (Fig. 1b), acting 3-5 h into the cell cycle in the G₁ phase, and controlling the entry into S phase (Fig. 2). Our results indicate that the murine counterpart of the human C3d-specific complement receptor CR2^{16,17} (known also to bind Sepharose-C3b) could be the signal-receiving receptor for the α -factors that directly or indirectly mimic the action of C3d. Human C3 fragments are active on murine cells. This predicts that the corresponding murine fragment, which is not available to us for experimentation, should be at least equally active.

We then tested the action of soluble C3d within the cell cycle of activated, synchronized B cells. Soluble C3d at 0.002-20 μ g ml⁻¹ did not replace α -factors in the stimulation of entry into S phase of activated B cells (tested in the presence of β -factors or of Sepharose-bound immunoglobulin-specific antibodies plus β -factors in conditions described in Fig. 1 legend, data not shown). In fact, soluble C3d, in concentrations ≥ 2 μ g ml⁻¹, completely inhibited the subsequent cell divisions that usually occur in the presence of Sepharose-bound immunoglobulin-specific antibodies and α - and β -factors (Fig. 3). These results suggest that the interaction of C3d-specific complement receptor CR2 with soluble C3d inhibits the stimulating action of α -factors, while crosslinking of the same receptors by Sepharose-bound C3d replaces the stimulating action of α -factors.

Our experiments identify C3d as a molecule that controls the entry of activated B cells into S phase. In this it resembles the

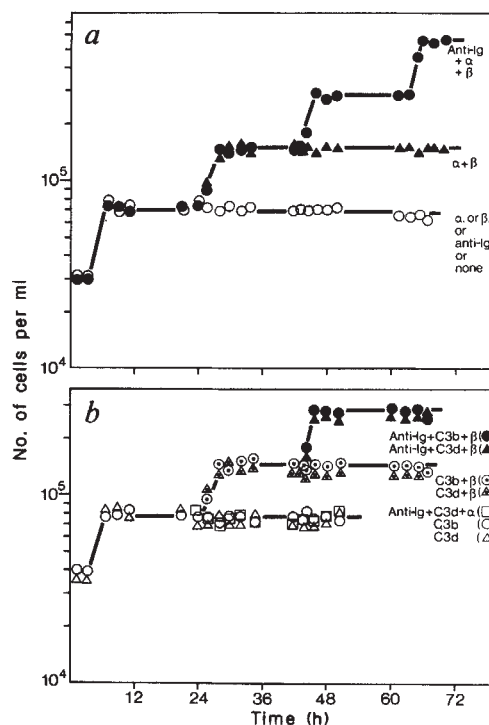
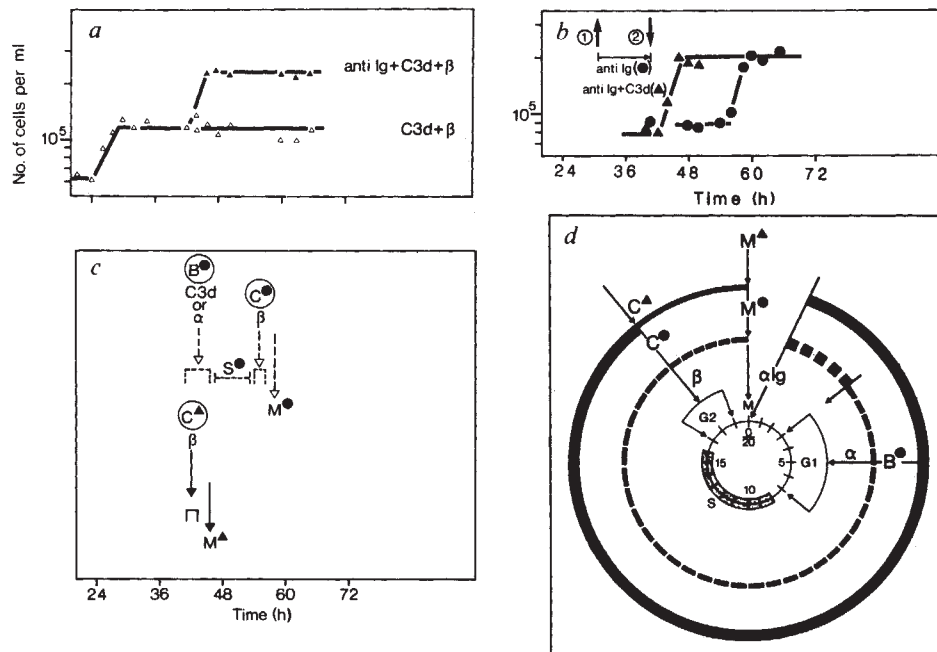


Fig. 1 Growth of B cells of C57BL/6J *nu/nu* mice, LPS activated and synchronized as described previously¹, in serum-substituted tissue cultures⁴⁰ containing either no additives, only α -factors, only β -factors, or only Sepharose-bound immunoglobulin-specific antibodies (reagents and conditions described in ref. 1). **a**, ○, No additives; ▲, cultures containing α - and β -factors; ●, cultures containing Sepharose-bound immunoglobulin-specific antibodies plus α - and β -factors. **b**, C3b-/C3d-dependent growth curves: in parallel cultures α -factors were replaced by Sepharose-bound C3b (○, △, ●) or C3d (△, △, ▲). C3 was purified as described elsewhere⁴¹, and converted to C3b by trypsin cleavage. C3d was purified from aged human plasma by sequential chromatography on DEAE-Sephadex and G100-Sephadex. It was >95% pure as judged by the silver-stained bands of C3d and C3d_g, detected in SDS-polyacrylamide electrophoresis. C3b and C3d were coupled to Sepharose 4B at respective concentrations of ~2 mg ml⁻¹ and 400 μ g ml⁻¹. Sepharose-bound C3b and C3d were also used as potential β -factor-replacing agents in cultures with anti-immunoglobulin plus α (shown for Sepharose-bound C3d, □). Since no divisions were observed with either C3b or C3d (shown only for C3d), we conclude that Sepharose-bound C3d did not replace β -factors. Cell concentrations in culture were determined by counting using a Bürker haemocytometer. Several counts were made with one 200- μ l culture. At least 50-100 cells were counted. With 100 cells counted, the accuracy of counting was determined in multiple samplings of the same culture to vary by <20%.

action of the family of insulin-like growth factors such as IGF-I, IGF-II, epidermal growth factor (EGF), nerve growth factor (NGF) and other factors that act on the cell cycle of eukaryotic cells of other differentiation lineages²¹⁻²⁵. We therefore sought for structural similarities between these insulin-like growth factors and C3d. The structure of the messenger RNA encoding the α -chain of C3, containing the proteolytic fragment C3d, has recently been reported by Wetsel *et al.*²⁶. A computer search identified a stretch of ~150 nucleotides between nucleotides 1,100 and 1,250 of the C3 α -chain mRNA (that is, within C3d) that was 50% homologous with a domain in the precursor mRNA of murine NGF²⁷ located between nucleotides 923 and 1,087, 45% homologous in the same stretch of sequence with murine IGF-II mRNA²⁸ between nucleotides 1 and 235 and 40% homologous with the precursor mRNA of murine EGF^{29,30} between 1,105 and 1,255. Furthermore, the computer identified two *Alu*-like sequences^{31,32} between nucleotides 1,501 and 1,560 with nucleotides 1,681 and 1,740 of the C3 α -chain mRNA, and *Alu*-like sequences of human ϵ -globulin³³ and human ALU7

Fig. 2 Action of Sepharose-bound immunoglobulin-specific antibodies (anti-Ig) and of Sepharose-bound C3d (C3d) in the cell cycle of activated, synchronized B cells (see Fig. 1). Cells described in Fig. 1b, which had divided after ~26 h in cultures containing Sepharose-bound C3d and β -factors (Δ), control cultures also containing Sepharose-bound, immunoglobulin-specific antibodies ($\blacktriangle$) (also shown in *a*) were washed free of C3d and β -factors at 30 h of culture (①↑) and resuspended with either anti-Ig ($\bullet$) or anti-Ig+C3d ($\blacktriangle$) between 30 and 40 h of culture (excitation phase). Anti-Ig, respectively anti-Ig+C3d, were removed (resuspending the cells by repeated pipetting, thereafter separating the Sepharose beads from the cells by differential sedimentation) and the bead-free cells were finally resuspended in cultures containing C3d+ β -factors (②↓). Growth of cells was monitored by cell counting (*b*). Cells incubated in the excitation phase with anti-Ig+C3d ($\blacktriangle$) divided 2–4 h later, that is, ~44 h of culture, while those incubated with anti-Ig alone ($\bullet$) divided 12–14 h later, ~58 h of culture. The model of the cell cycle of B cells (*c*, *d*) shows that anti-Ig acts directly after mitosis (M) at a first restriction point (A) and moves the cells to a second restriction point (B), between 4 and 8 h after M, where α -factors act. β -Factors act at a third restriction point (C), 2–4 h before M, after S phase (DNA replication). C3d replaces α -factor action. Its absence during the excitation phase arrests cells 12–14 h before M, like α -factors ($\bullet$, dashed lines; in *d* the thickly drawn line), before S, while its presence moves the cells to the restriction point for β , 2–4 h before M ($\blacktriangle$, solid lines; in *d* the thickly drawn part). In *c* the phases of the B-cell cycle which are completed to reach mitosis (M) in the presence of C3d and β -factors after excitation with anti-Ig alone (dashed lines) or anti-Ig plus C3d (solid lines). In *d*, excitation phases (drawn thickly) in the presence of anti-Ig (inner, dashed circle) or anti-Ig plus C3d (outer, solid circle), and post-excitation phases in the presence of C3d and β -factors are shown within the cell cycle relative to mitosis (M).



(ref. 34), respectively, flanking a sequence of 44% homology between nucleotides 1,560 and 1,650 of the C3 α -chain mRNA and nucleotides 800 and 950 of the murine oncogene *c-mos* (refs 35, 36).

Although the sequence homologies of C3d with the insulin-like growth factor precursors are weak, they are intriguing

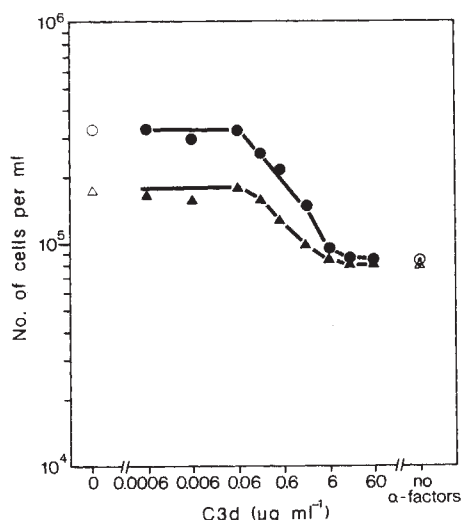


Fig. 3 Influence of soluble C3d on the growth of activated, synchronized C57BL/6J *nu/nu* splenic B cells cultured in the presence of Sepharose-bound immunoglobulin specific antibodies, α -factors and β -factors (see in Fig. 1a). Different concentrations of soluble C3d were added to the cultures at 10 h of incubation (see Fig. 1a). Cell concentrations were determined at 32 h (Δ), that is, after one division, and at 50 h ($\bullet$), after two divisions. In the presence of Sepharose-bound immunoglobulin-specific antibodies, α - and β -factors, but no C3d, the cells divided (Δ , $\circ$), whereas the omission of α -factors ('No α -factors') abolished the capacity of the activated B cells to divide (Δ , $\circ$).

because NGF, IGF-II and EGF are similar in size (50–70 amino acids) and function (in their control of entry into S phase). It will be interesting to investigate whether C3d is, in fact, further processed by proteolytic cleavage to yield such an ~50–70-amino-acid polypeptide with α -factor activity. Furthermore, this sequence similarity places C3 in phylogenetic relation to the blood clotting factors IX and X and to bovine protein C (ref. 37). Because of the homology that was recently found between the receptors for EGF³⁸ and insulin³⁹, it will be interesting to see whether such a phylogenetic relationship exists between these receptors and the C3d-specific complement receptor CR2. The significance of the structural homologies of C3d to *Alu*-like sequences and to *c-mos* will remain unclear until the normal function of the oncogene is known.

We thank Drs R. Palacios and C. Paige for critical reading of our manuscript. The Basel Institute for Immunology was founded and is entirely supported by F. Hoffmann-La Roche, Ltd., Co., Basel, Switzerland. A.E. was on leave of absence from the Department of Immunology, L. Eötvös University, Göd, Hungary. T.S. and M.P.D. were supported by grant P5253 from the Fonds zur Förderung der wissenschaftlichen Forschung.

Received 23 April; accepted 24 July 1985.

- Melchers, F., Corbel, C., Leptin, M. & Lernhardt, W. *J. Cell Sci.* (in the press).
- Hoffmann, M. K., Hämmerling, U., Simon, M. & Oettgen, H. F. *J. Immun.* **116**, 1447–1451 (1976).
- Corbel, C. & Melchers, F. *Immun. Rev.* **78**, 51–74 (1984).
- Corbel, C. & Melchers, F. *Eur. J. Immun.* **13**, 528–533 (1983).
- Andersson, J. & Melchers, F. *Cell* **37**, 715–720 (1984).
- Pepys, M. B. *Nature new Biol.* **237**, 157–159 (1972).
- Dukor, P. & Hartmann, K.-U. *Cell. Immun.* **7**, 359–356 (1973).
- Hartmann, K.-U. & Bokisch, V. A. *J. exp. Med.* **142**, 600–610 (1975).
- Bitter-Suermann, D. *et al. J. Immun.* **111**, 301 (1973).
- Dukor, P. *et al. J. exp. Med.* **139**, 337–354 (1974).
- Pepys, M. B. *J. exp. Med.* **140**, 126–144 (1974).
- Klaus, G. G. B. & Humphrey, J. H. *Immunology* **33**, 31–40 (1977).
- Lai, A., Fat, R. F. M. & Van Furth, R. *Immunology* **28**, 359–368 (1975).
- Whaley, K. J. *J. exp. Med.* **151**, 501–516 (1980).
- Erdei, A., Melchers, F., Schulz, T. & Dierich, M. *Eur. J. Immun.* **15**, 184–188 (1985).
- Ross, G. D. *Fedn Proc.* **41**, 3089–3093 (1982).
- Fearon, D. T. *Immun. Today* **5**, 105–110 (1984).
- Miller, R. G. & Phillips, R. A. *J. cell Physiol.* **73**, 191–202 (1969).

19. Andersson, J., Sjöberg, O. & Möller, G. *Transplant Rev.* **4**, 131-177 (1972).
20. Andersson, J., Coutinho, A. & Melchers, F. *J. exp. Med.* **149**, 553-564 (1979).
21. Lindgren, A. & Westermark, B. *Expl Cell Res.* **99**, 357-362 (1976).
22. Lindgren, A., Westermark, B. & Pontén, J. *Expl Cell Res.* **95**, 311-319 (1975).
23. Stiles, C. D. *et al. Proc. natn. Acad. Sci. U.S.A.* **76**, 1279-1283 (1979).
24. Leof, E. B., Wharton, W., van Wyk, J. J. & Pledger, W. J. *Expl Cell Res.* **141**, 107-115 (1982).
25. Wharton, W., van Wyk, J. J. & Pledger, W. J. *J. cell. Physiol.* **107**, 31-39 (1981).
26. Wetsel, R. A. *et al. J. Biol. Sci.* **259**, 13857-13862 (1984).
27. GenBank, genetic bank sequence, locus MUSNGF, Release 26.0 (1984).
28. Bell, G. I. *et al. Nature* **310**, 775-777 (1984).
29. Gray, A., Dull, T. J. & Ullrich, A. *Nature* **303**, 722-725 (1983).
30. Scott, J. *et al. Science* **221**, 236-240 (1983).
31. Jelinek, W. R. & Schmid, C. W. *Rev. Biochem.* **51**, 813-844 (1982).
32. Ullu, E. & Tschudi, Ch. *Nature* **312**, 171-172 (1984).
33. Gen Bank, genetic bank sequence, locus NUMHBE Release 26.0 (1984).
34. Gen Bank, genetic bank sequence, locus HUMALU7, Release 26.0 (1984).
35. Van Beveren, C. *et al. Nature* **289**, 258-262 (1981).
36. Van Beveren, C., van Straaten, F., Galleshaw, J. A. & Verma, I. M. *Cell* **27**, 97-108 (1981).
37. Doolittle, R. F., Feng, D. F. & Johnson, M. S. *Nature* **307**, 558-560 (1984).
38. Ullrich, A. *et al. Nature* **309**, 418-425 (1984).
39. Ullrich, A. *et al. Nature* **313**, 756-761 (1985).
40. Iscove, N. N. & Melchers, F. *J. exp. Med.* **147**, 923-933 (1978).
41. Hammer, C. H., Wirtz, G. H., Reufer, L., Gresham, H. D. & Tack, B. F. *J. Biol. Chem.* **256**, 3995-4003 (1981).

Sequence of protein disulphide isomerase and implications of its relationship to thioredoxin

Jeffrey C. Edman*, Leland Ellis*, Russell W. Blachert†, Richard A. Horst‡ & William J. Rutter*

* Hormone Research Laboratory, University of California, San Francisco, California 94143, USA

† Chiron Corporation, 4560 Horton St, Emeryville, California 94608, USA

‡ Department of Pharmacology, Stanford University School of Medicine, Stanford, California 94305, USA

The formation of disulphide bonds is essential to the structure and function of proteins. These bonds rapidly form either co-translationally or immediately post-translationally in the lumen of the endoplasmic reticulum^{1,2}. Native disulphide pairing for such proteins has been achieved *in vitro*; however, the rates of reassembly are slow and the conditions non-physiological^{3,4}. To account for these observations, Anfinsen *et al.* proposed that a 'disulphide interchange protein' was the *in vivo* catalyst of disulphide bond rearrangement⁵. Other groups discovered an activity with similar characteristics that catalysed the reductive cleavage of insulin and may be associated with insulin degradation^{6,7}, although this result has been disputed⁸. The enzyme involved, protein disulphide isomerase (PDI; EC 5.3.4.1), may be the *in vivo* catalyst of disulphide bond formation⁹. Here we describe the sequence of cloned rat liver PDI complementary DNA which predicts a protein with two distinct regions homologous with *Escherichia coli* thioredoxin, a known cofactor in oxidation-reduction reactions¹⁰. Each of these regions contains the presumed active site sequence Trp-Cys-Gly-His-Cys-Lys, suggesting that PDI, similar in action to thioredoxin, catalyses disulphide bond interchange via an internal disulphide-sulphydryl interchange. The cDNA predicts a signal peptide consistent with the view that PDI is a luminal endoplasmic reticulum protein. PDI messenger RNA, although ubiquitous, is more highly concentrated in secretory cells.

PDI catalyses the isomerization of both intramolecular^{11,12} and intermolecular (for example, IgM; see ref. 13) disulphide bonds yielding apparently native structures. Despite these observations, there is still uncertainty about the physiological role of PDI. Many proteins can form proper disulphide bonds in the absence of PDI *in vitro*¹⁴, and the level of stimulation with added PDI in some cases is only a few times the background¹⁵. Furthermore, nearly stoichiometric quantities of PDI are often necessary for optimal reaction rates. To address the requirements of disulphide bond formation *in vivo*, we have cloned sequences of rat liver and pancreas PDI cDNA, determined its primary sequence and analysed the tissue distribution of PDI mRNA.

A rat pancreas cDNA library cloned into the expression vector λ gt11 (ref. 16) was screened using a monoclonal antibody to rat liver PDI¹⁷. From 80,000 recombinant phage, a single clone (λ pdi1) reactive to the antibody was isolated. This clone has a 1,100 base pair (bp) insert containing approximately half of the predicted coding sequence for PDI. Rescreening of the pancreas library with ³²P-labelled λ pdi1 insert probe revealed no phage with more 5' sequences. Screening of a rat liver cDNA library in λ gt10 (ref. 18), however, allowed us to isolate λ pdi6 (insert size 2,400 bp), which was subcloned (ppdi6) and sequenced (see Fig. 1). By primer extension analysis (data not shown), λ pdi6 was lacking ~160 bp on its 5' end. A plasmid clone (ppdi306) containing the 5'-most sequences was isolated as described in Fig. 1 legend. All sequence information presented here was determined from rat liver cDNAs. No differences in coding information were noted in comparing λ pdi1 (pancreas) with liver sequences; however, they do use alternate polyadenylation sites (see below).

The identity of λ pdi6 was confirmed using hybrid-arrested translation¹⁹ of *in vitro* synthesized and immunoprecipitated PDI. Affinity-purified rabbit anti-PDI antibodies precipitate a protein of relative molecular mass (M_r) 57,000 (57K) from the total *in vitro* translation products of liver and pancreas poly(A)⁺ RNA (Fig. 2A, c, e). In parallel experiments, hybridization of total pancreas poly(A)⁺ RNA to denatured ppdi6 completely abolished synthesis of PDI (Fig. 2B, a), whereas heat treatment of the hybrid released PDI mRNA for translation (b). Further confirmation was obtained by comparing the N-terminal sequence of highly purified rat liver PDI¹⁷ with the sequence predicted from the cDNA. The sequence obtained from the PDI protein was (Trp?)-Ala-Leu-Glu-Glu-Glu-Asp-Asn-Val-Leu-Val-Leu. The first cycle (Trp?) did not agree with the predicted sequence (Asp), but was obtained in low yield. The remaining sequence precisely matches the predicted sequence of the cDNA and is underscored in Fig. 1B.

PDI mRNA contains ~40 bases of 5' untranslated sequence and a coding region of 1,524 bases. At least two polyadenylation sites are used. The first of these lies 221 bases downstream from the termination codon and is not preceded by the canonical AAUAAA hexanucleotide. This polyadenylation site was used in the first clone isolated (λ pdi1), but was not found in two other pancreas cDNA clones. The second lies 891 bases downstream and has the AAUAAA sequence 20 bases 5' to the poly(A) tract. The existence of such multiple polyadenylation sites is well known in other gene systems, but their function is unclear²⁰. By Northern analysis (Fig. 3A), a major species of PDI mRNA at 2,800 bases and a very minor component of 2,000 bases were detected, corresponding to the polyadenylation sites mentioned above. The distribution of PDI mRNA was investigated by measurement with labelled insert probe and RNA dot blots (see Fig. 3B). The relative content of PDI mRNA is liver > pancreas ~ kidney > lung > testes ~ spleen > heart > brain. This series is consistent with the known tissue distribution of PDI activity³. By cloning frequency, PDI mRNA comprises 0.01-0.05% of total poly(A)⁺ from rat liver. The apparent discrepancy between mRNA level and protein content (0.4%) is probably a result of the long half-life of PDI (> 7 days; see ref. 21).

The cloned sequences encode a protein of 508 amino acids (M_r 56,783) in good agreement with the molecular size determination of *in vitro*-translated PDI. Preceding the sequence obtained by N-terminal sequencing are 19 predominantly hydrophobic amino acids, characteristic of a signal peptide²². The presence of a signal peptide strongly suggests that PDI is transported into the lumen of the endoplasmic reticulum²³. The predicted amino-acid composition is in agreement with that determined from rat liver PDI (data not shown, ref. 24). No N-linked glycosylation sites (Asn-X-Ser/Thr) are present; therefore, if PDI is glycosylated, only O-linked sugars are present. The sequence predicts an acidic molecule with a pI of 4.4. A pI of 4.2-4.3 has been detected for PDI isolated from rat liver²⁵.

The mature protein contains two pairs of regions with internal homology. Using these homologies as guideposts, we have



Fig. 2 *In vitro* translation and hybrid-arrested translation of PDI mRNA. **A**, *a-c*, Rat pancreas poly(A)⁺ RNA; *b*, normal rabbit IgG immunoprecipitate; *c*, rabbit anti-PDI immunoprecipitate; *d*, normal rabbit IgG immunoprecipitate; *e*, rabbit anti-PDI immunoprecipitate; *f*, total translation products of liver poly(A)⁺ RNA; *g*, minus-added RNA control. **B**, Hybrid-arrested translation of PDI mRNA from rat pancreas poly(A)⁺ RNA. *a*, Hybridization of poly(A)⁺ RNA with p_{pd16}; *b*, hybridization as in *a* except hybrids were treated at 100 °C for 1 min to release hybridized mRNA; *c*, the control hybridization of poly(A)⁺ RNA with pMT21; *d*, hybridization as in *c* except that hybrids were treated at 100 °C for 1 min to release hybridized mRNA. Numbers on the left refer to *M_r* standards ×1,000; phosphorylase, 92K; bovine serum albumin, 68K; ovalbumin, 43K; and carbonic anhydrase, 30K.

Methods. Poly(A)⁺ RNA (5–10 µg) from adult rat pancreas or liver was translated *in vitro* using a rabbit reticulocyte lysate (Promega Biotec) according to the manufacturer's instructions in the presence of ³⁵S-methionine (NEN). PDI was immunoprecipitated from the total translation products using affinity-purified rabbit anti-rat PDI (1.5–5.0 µg ml⁻¹) and *Staphylococcus aureus* cells (Pansorbin, Calbiochem)⁴⁰. Control precipitations used normal rabbit IgG (1.5–5.0 µg ml⁻¹; Cappel). Hybridization-arrested translation was performed as described¹⁹ using 2 µg linearized plasmid DNAs. Samples were run on 5% polyacrylamide gels⁴¹; dried gels were used for autoradiography at -70 °C with X-Omat film (Kodak).

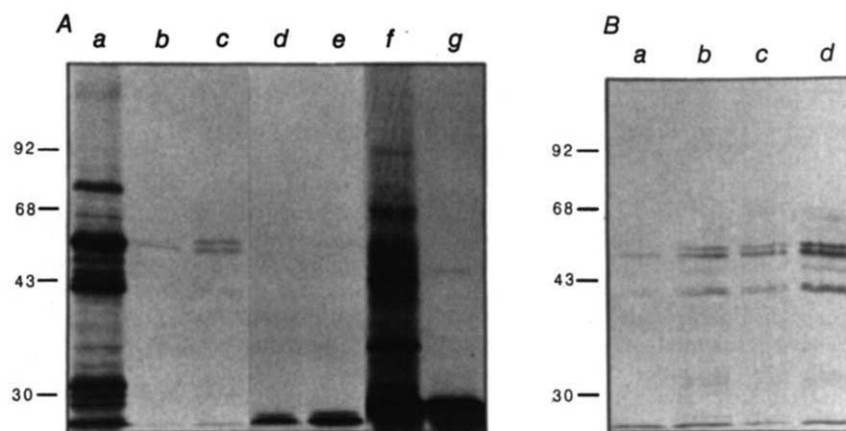
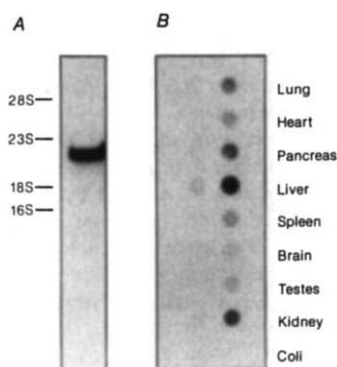


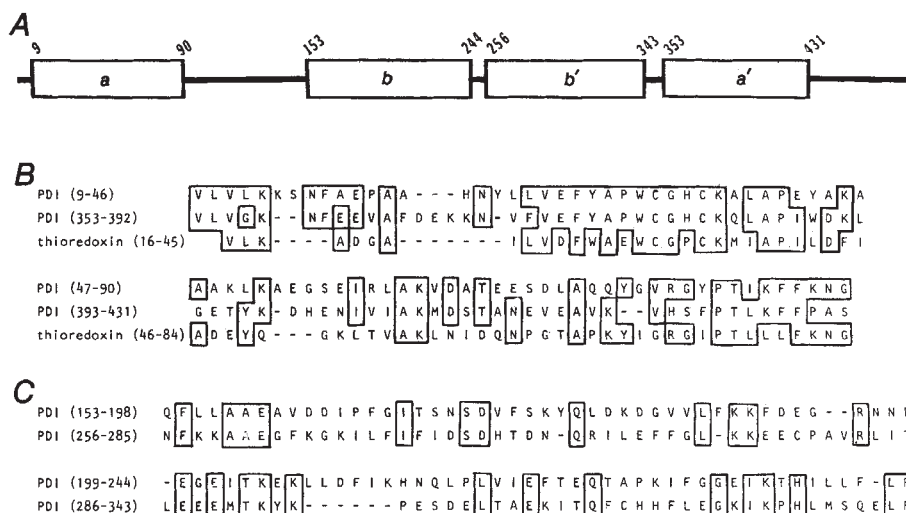
Fig. 3 Size analysis and distribution of PDI mRNA. **A**, Northern gel of poly(A)⁺ RNA from rat pancreas.

2 µg total poly(A)⁺ RNA from rat pancreas was electrophoresed on a 1% agarose/2.2 M formaldehyde gel and blotted to nitrocellulose according to Thomas⁴², and hybridized with ³²P-labelled λ_{pd16} insert probe. Size markers on the left are 28S and 18S from rat pancreas and 23S and 16S from *E. coli*. **B**, Tissue distribution of PDI mRNA. One µg and 10 µg of total RNA isolated from the tissues indicated and *E. coli*⁴³ was denatured with 2.2 M formaldehyde in 10 mM sodium phosphate, pH 7.0, for 10 min at 60 °C. RNAs were diluted 10-fold with ice-cold 15 × SSC and blotted to nitrocellulose using a manifold apparatus (Schleicher & Schuell). The filter was baked and hybridized to ³²P-labelled total insert probe derived from λ_{pd16}. Autoradiography was performed overnight at -70 °C with a Dupont Cronex lightning plus screen and X-Omat film (Kodak).



by the sulphhydryl reagent iodoacetate²⁷. The sequences around these Cys pairs also correspond well with the amino-acid compositions obtained by DeLorenzo *et al.*²⁸ of tryptic peptides from beef liver PDI labelled by ¹⁴C-iodoacetate. Downstream from these sequences the homology decreases, but basic and aromatic residues generally remain conserved. Comparison of regions *a* and *a'* with *E. coli* thioredoxin²⁹ revealed a considerable sequence homology (see Fig. 4B). Within the 16 amino-acid regions noted above, 10 amino acids are identical, 4 show functional conservation and 2 (His in PDI and Pro in thioredoxin; Pro in PDI and Glu in thioredoxin) are different. Secondary structure analysis³⁰ of PDI reveals that these regions, like thioredoxin, are probably situated on a reverse turn between a short β-stretch and a long α-helix (see Fig. 1B). Again the homology decreases beyond this highly conserved area, but specific conserved residues are still apparent. Regions *b* and *b'* situated between *a* and *a'* share no significant homology with any other known proteins. Despite considerable divergence in primary structure between *b* and *b'*, predicted secondary structure is remarkably similar (see Fig. 1B). In addition, the distribution of charged and hydrophilic residues is also similar. The two remaining regions of PDI (amino acids 91–152 and 432–489) also do not demonstrate any significant homology with known proteins. The region 91–152 is predicted to be almost entirely helical and disrupted by two reverse turns. The region 431–489 contains a high density of negatively charged residues (29% Glu or Asp).

Fig. 4 **A**, Schematic illustration of internal homologies within PDI. Numbers above refer to residue numbers (see text). **B**, Alignment of homologies with PDI regions *a* and *a'* and with *E. coli* thioredoxin. Residue numbers are shown on the left in parentheses. Identical residues are boxed. For clarity, conservative replacements have not been shown. **C**, Homology between regions *b* and *b'* within PDI. The sequences are shown in single-letter codes that correspond to the three-letter codes as follows: A, Ala; C, Cys; D, Asp; E, Glu; F, Phe; G, Gly; H, His; I, Ile; K, Lys; L, Leu; M, Met; N, Asn; P, Pro; Q, Gln; R, Arg; S, Ser; T, Thr; V, Val; W, Trp; Y, Tyr.



The two Cys residues of thioredoxin function as an oxidation-reduction couple²⁹. Thioredoxin can catalyse the reduction of insulin, one of the known activities of PDI³¹. Analogous to this action of thioredoxin is the function of the two pairs of PDI cysteines, which cooperate in the formation of disulphide bonds. The formation of a disulphide bond in a nascent polypeptide chain requires the presence of an oxidizing equivalent and juxtaposition of the two Cys residues involved. Creighton *et al.*¹² proposed a mechanism by which the reduced protein first reacts with oxidized PDI and forms a mixed disulphide, which then undergoes an intramolecular disulphide rearrangement to yield a disulphide bond and reduced PDI. Oxidized PDI would then be regenerated by a coupled oxidizing system (for example, oxidized glutathione). Non-native disulphide bonds formed as intermediates in the folding pathway^{4,12} could be degraded by the reverse of this process. This process would continue until the final structure is obtained and disulphides would presumably no longer react because of thermodynamic constraints and/or physical inaccessibility. The presence of a signal peptide suggests that PDI exists within the endoplasmic reticulum; as PDI associates with itself⁹, a nested array of active-site disulphides may present the nascent polypeptide chain with an oxidizing microenvironment that allows the efficient catalysis of disulphide interchange.

Our finding of two domains containing active sites suggests an interesting alternative: because each domain can ostensibly react with sulphhydryl groups or disulphide bonds, the condition for appropriate disulphide bond formation might be that the intrinsic folding energy of the interacting protein brings the two PDI active sites in apposition, thus facilitating disulphide bond interchange. The availability of the cDNA should allow direct analysis of the role of PDI in the process of disulphide bond interchange.

We thank Dr Robert Stroud for secondary structure prediction and homology alignments. We also thank Dr Frank Masiarz of Chiron Corporation for the protein sequence facility, Jennifer Turner and Jennifer Barnett for DNA preparations and oligonucleotides, Drs James Ou and Charles Craik for guidance on sequencing, Leslie Spector for manuscript preparation and our colleagues in the Hormone Research Institute for their support. This work was supported by NIH grants AM21344 to W.J.R. and AM34926 to R.A.R. L.E. received a Senior Fellowship from the California Division of the American Cancer Society.

Received 2 April; accepted 25 June 1985.

- Bergman, L. W. & Kuehl, W. M. *J. biol. Chem.* **254**, 5690-5694 (1979).
- Peters, T. & Davidson, L. K. *J. biol. Chem.* **257**, 8847-8853 (1982).
- Freedman, R. B. & Hillson, D. A. in *The Enzymology of Post-Translational Modifications of Proteins* (eds Freeman, R. B. & Hawkins, H. C.) 167-212 (Academic, New York, 1980).
- Anfinson, C. B. & Scheraga, H. A. *Adv. Protein Chem.* **29**, 205-300 (1975).
- Goldberger, R. F., Epstein, C. J. & Anfinsen, C. B. *J. biol. Chem.* **238**, 628-635 (1963).
- Tomizawa, H. *J. biol. Chem.* **237**, 428-431 (1962).
- Katzen, H. M. & Stetten, D. *Diabetes* **11**, 271-280 (1962).
- Duckworth, W. C. & Kitabchi, A. E. *Endocrine Rev.* **2**, 778-782 (1983).
- Freedman, R. B. *Trends biochem. Sci.* **9** (1984).
- Holmgren, A. *Eur. J. Biochem.* **6**, 475-484 (1968).
- Freedman, R. B., Brockway, B. E. & Lambert, N. *Biochem. Soc. Trans.* **12**, 929-932 (1984).
- Creighton, T. E., Hillson, D. A. & Freedman, R. B. *J. molec. Biol.* **142**, 43-62 (1980).
- Roth, R. A. & Koshland, M. E. *Biochemistry* **20**, 6594-6599 (1981).
- Wetlauffer, D. B., Saxena, V. P., Ahmed, A. K., Schaffer, S. W., Pick, P. W., Oh, K.-J. & Peterson, J. D. in *Protein Crosslinking Pt A* (ed. Friedman, M.) 43-50 (Plenum, New York, 1977).
- Teale, J. M. & Benjamin, D. C. *J. biol. Chem.* **251**, 4603-4608 (1976).
- Young, R. A. & Davis, R. W. *Science* **222**, 778-782 (1983).
- Roth, R. A. & Mesirov, M. L. *Biochim. biophys. Acta* **788**, 189-192 (1984).
- Hyunh, T., Young, R. & Davis, R. W. in *Practical Approaches in Biochemistry* (ed. Glover, D.) (IRL, Oxford, in the press).
- Paterson, B. M., Roberts, B. E. & Kuff, E. L. *Proc. natn. Acad. Sci. U.S.A.* **74**, 4370-4374 (1977).
- Reynolds, G. A. *et al. Cell* **38**, 275-285 (1984).
- Ohba, H., Harano, T. & Omura, T. *J. Biochem., Tokyo* **213**, 225-234 (1983).
- Sabatini, D. D., Kreibich, G., Morimoto, T. & Adesnik, M. *J. Cell Biol.* **92**, 1-22 (1982).
- Okada, Y. *et al. Eur. J. Biochem.* **122**, 393-402 (1982).
- Ansorge, S. *et al. FEBS Lett.* **37**, 238-240 (1973).
- Mills, E. N. C., Lambert, N. & Freedman, R. B. *Biochem. J.* **213**, 245-248 (1983).
- Hillson, D. A. & Freedman, R. B. *Biochem. Soc. Trans.* **7**, 573-574 (1979).
- Carmichael, D. F., Keefe, M., Pace, M. & Dixon, J. E. *J. biol. Chem.* **254**, 8386-8390 (1979).
- DeLorenzo, F., Fuchs, S. & Anfinsen, C. B. *Biochemistry* **5**, 3961-3965 (1966).
- Holmgren, A., Soderberg, B. O., Eklund, H. & Branden, C. I. *Proc. natn. Acad. Sci. U.S.A.* **72**, 2305-2309 (1975).
- Finer-Moore, J. & Stroud, R. M. *Proc. natn. Acad. Sci. U.S.A.* **81**, 155-159 (1984).
- Holmgren, A. *J. biol. Chem.* **254**, 9627-9632 (1979).
- Maniatis, T., Fritsch, E. F. & Sambrook, J. *Molecular Cloning, A Laboratory Manual* (Cold Spring Harbor Laboratory, New York, 1982).

- Aviv, H. & Leder, P. *Proc. natn. Acad. Sci. U.S.A.* **69**, 1408-1412 (1972).
- Gubler, U. & Hoffman, B. J. *Gene* **25**, 263-269 (1983).
- Hanahan, D. *J. molec. Biol.* **166**, 557-580 (1983).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Hunkapiller, M. W., Hewich, R. M., Dreyer, W. J. & Hood, L. E. *Meth. Enzym.* **91**, 399 (1983).
- Hawke, D., Yuan, P. M. & Shively, J. E. *Analyt. Biochem.* **120**, 302 (1982).
- Arias, C. F., Lopez, S., Bell, J. R. & Strauss, J. H. *J. Virol.* **50**, 657-661 (1984).
- Kessler, S. W. *J. Immun.* **115**, 1647-1624 (1975).
- Laemmli, U. K. *Nature* **227**, 680-685 (1970).
- Thomas, P. S. *Meth. Enzym.* **100**, 255-266 (1983).
- Chirgwin, J. M., Przybyla, A. E., MacDonald, R. J. & Rutter, W. J. *Biochemistry* **18**, 5294-5301 (1979).

Primary structure of the precursor to the three major surface antigens of *Plasmodium falciparum* merozoites

Anthony A. Holder, Michael J. Lockyer, Karel G. Odink*, Jasbir S. Sandhu, Valentina Riveros-Moreno, Stephen C. Nicholls, Yvonne Hillman, Lynne S. Davey, Mark L. V. Tizard, Ralph T. Schwarz & Robert R. Freeman

Department of Molecular Biology, Wellcome Biotechnology Ltd, Langley Court, Beckenham BR3 3BS, UK

Recently, a class of protein antigens of high relative molecular mass (M_r) which can induce protective immunity against blood-stage malaria has been identified¹⁻⁵. In *Plasmodium falciparum* the protein has a M_r of ~195,000 (P195)^{6,7}. It is the precursor of three proteins of M_r 83,000 (83K), 42K and 19K which are the major surface antigens of merozoites^{8,9}; thus it may also be useful for immunization against *P. falciparum*. Three studies describing the isolation of single short complementary DNA clones for part of the P195 gene sequence have been reported^{4,10,18}. Here we describe the complete structure of the P195 gene determined from further DNA clones, its organization within genomic DNA and the location of the specific processing fragments within the primary amino-acid sequence.

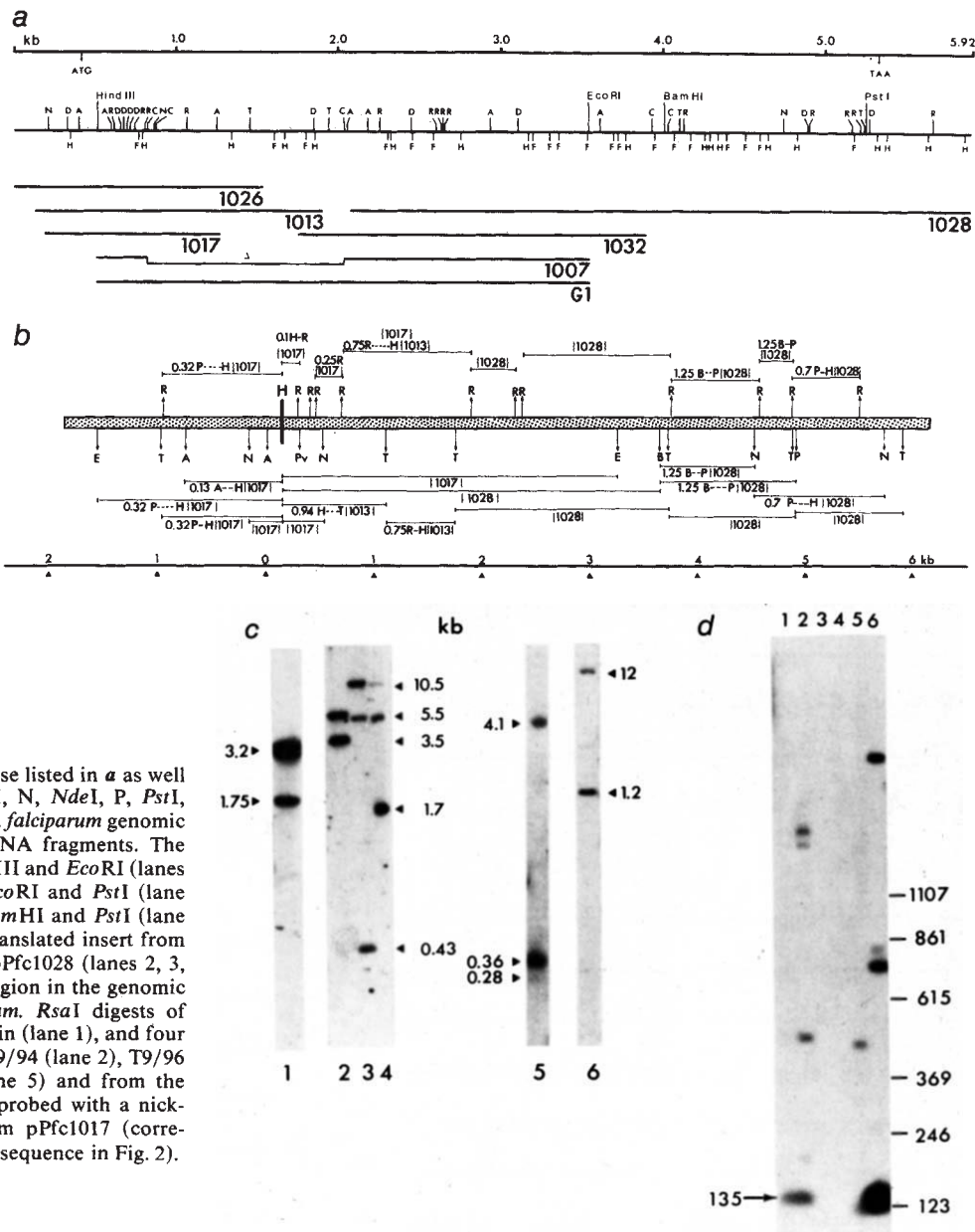
The cDNA clone previously isolated¹⁰ was used as a probe to screen cDNA and genomic recombinants to obtain further DNA clones that cover the complete coding sequence (Fig. 1a). A total of 45 cDNA clones and 1 genomic clone from a *Hind*III-*Eco*RI digest of genomic DNA were isolated and characterized by cross-hybridization and restriction enzyme mapping. The cDNA clones and specific fragments from them were used to probe Southern blots of *P. falciparum* DNA cut with restriction enzymes (Fig. 1c). The map of restriction sites in genomic DNA (Fig. 1b) was colinear with that derived from the cDNA nucleotide sequence over the region corresponding to the coding sequence of the protein. This indicates that there are no discontinuities (introns) in the genomic DNA compared with the messenger RNA sequence.

The sequence of 5,920 nucleotides is shown in Fig. 2. A methionine start codon at nucleotide 418 is followed by an open reading frame of 4,920 nucleotides until the first stop codon at position 5,337. The A+T content of the sequence is high, an average of 76% within the coding region and 90% within the untranslated 5' and 3' ends. The open reading frame codes for a protein containing 1,640 amino acids with a calculated M_r of 188,049. The start codon is followed by a sequence coding for a putative signal peptide of 18 amino acids and there is a hydrophobic tail sequence at the carboxy terminus. Within the amino-acid sequence at residues 72-104, an alternate repeat of the two tripeptide sequences Ser-Gly-Gly and Ser-Val-Ala occur six and five times respectively. At the nucleotide level the central core of this repeat consists of an 18-base-pair (bp) sequence repeated four times with no degeneracy in the codon usage.

Of the 19 cysteine residues, 11 are in the 96 C-terminal amino

* Present address: Ciba-Geigy AG, K-681-442, 4002 Basel, Switzerland.

Fig. 1 a, Restriction map of the P195 gene and alignment of the DNA clones used for sequence analysis. The clone pPfc1007 contained an internal deletion as indicated. The restriction enzyme sites indicated are: A, *AluI*; C, *HincII*; D, *DdeI*; F, *HinfI*; H, *AhaIII*; N, *NdeI*; R, *RsaI*; T, *TaqI*. The positions of the start (ATG) and stop (TAA) codons for the long open reading frame are indicated on the scale bar. Complementary DNA from *P. falciparum* polyadenylated RNA and cDNA libraries was prepared as described previously¹⁰. Dideoxy chain termination and chemical cleavage^{12,13} were used to determine the sequence. Storage and manipulation of the sequence data were performed on a VAX 11/750 computer using the programs of Staden^{14,15}. **b**, Organization of the gene in genomic DNA from the Wellcome strain. The bars above and below the gene indicate the labelled probe used to identify specific restriction fragments. When specific fragments of cDNA clones were used, they are presented as, for instance, 0.32P-H(1017), meaning the 0.32-kilobase (kb) *PstI*-*HindIII* fragment from pPfc1017. The restriction enzyme sites indicated include those listed in **a** as well as: B, *BamHI*; E, *EcoRI*; H, *HindIII*; N, *NdeI*; P, *PstI*; Pv, *PvuII*. **c**, Southern blot analysis of *P. falciparum* genomic DNA probed with specific cloned DNA fragments. The genomic DNA was digested with *HindIII* and *EcoRI* (lanes 1, 2), *EcoRI* and *BamHI* (lane 3), *EcoRI* and *PstI* (lane 4), *HindIII* and *NdeI* (lane 5) and *BamHI* and *PstI* (lane 6). The blots were probed with nick-translated insert from clone pPfc1017 (lanes 1, 5) or clone pPfc1028 (lanes 2, 3, 4, 6). **d**, Presence of the P195 repeat region in the genomic DNA of cloned lines of *P. falciparum*. *RsaI* digests of genomic DNA from the Wellcome strain (lane 1), and four cloned lines from T9, a Thai isolate, T9/94 (lane 2), T9/96 (lane 3), T9/98 (lane 4), T9/101 (lane 5) and from the cDNA clone pPfc1013 (lane 6) were probed with a nick-translated 135-bp *RsaI* fragment from pPfc1017 (corresponding to nucleotides 613–747 in the sequence in Fig. 2).



acids. There are 12 potential *N*-glycosylation sites and the P195 of several strains has been shown to be glycosylated (R.T.S., unpublished observations). A comparison of the nucleotide sequence with that determined by Hall *et al.*⁴ for a cDNA clone of P190 derived from K1, a Thai strain of *P. falciparum*, shows that their cDNA sequence begins at position 970 and is identical with the sequence in Fig. 2 for 164 bases with the exception of an A–G substitution at position 973. The remainder of the sequence of Hall's clone is related to but not identical with the sequence up to nucleotide 1,248. Although the significance of the differences between the two sequences is unclear, the position of homology does not support their suggestion that this region is derived from the 3' end of mRNA.

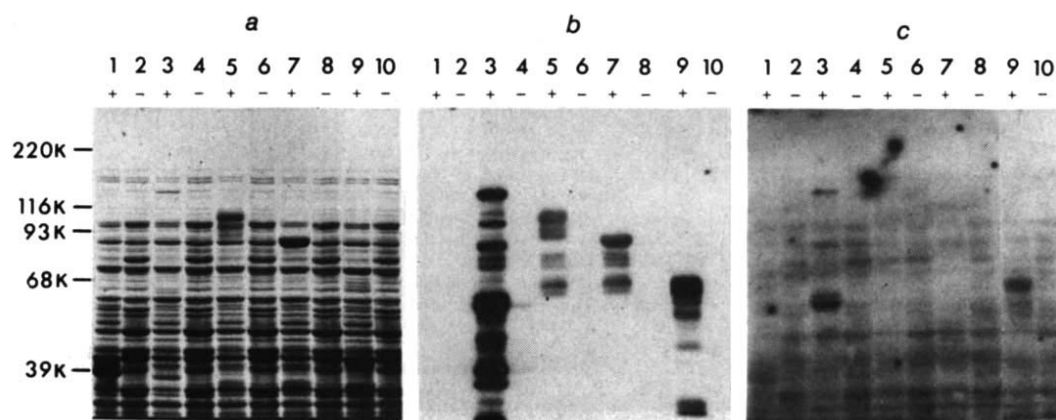
We looked for the presence of the P195 repeats within the DNA of four cloned lines of *P. falciparum* derived from T9, a Thai isolate¹¹. The *RsaI* fragment spanning the repeat sequence is present in the genomic blots of the Wellcome (West African) strain, in T9/94 and as a slightly faster moving fragment in T9/101, but is completely absent in T9/96 and T9/98 DNA (Fig. 1d). Therefore, this particular structure is variable in different clones from a single *P. falciparum* isolate. If the repeats are immunogenic they may represent a strain-specific repeating epitope and other clones may have a different repetitive structure.

Specific fragments of the cloned DNA were expressed in *Escherichia coli* as fusion proteins (Fig. 3). Inducible fusion proteins of *M_r* 135K, 105K, 85K and 65K were present in the lysates of strains containing the plasmids pME1, pME2, pME3 and pME4, respectively, detected by Coomassie blue staining or Western blotting with polyclonal anti-P195 antiserum. Additional lower-*M_r* species reactive with the antibodies that were present could be due to proteolysis or premature termination of the fusion proteins.

The position of the processed fragments in the protein sequence was determined by direct sequencing or by binding to specific antibodies. The 83K fragment from P195 was purified in a soluble form (Fig. 4a) and subjected to N-terminal sequence analysis. The amino acids that were identified, positions 6–20 from the N-terminus, correspond to residues 25–39 in the sequence (Fig. 2). This indicates that the 83K fragment is located at the amino terminus of P195 and that cleavage of the putative signal peptide occurs between the residues at positions 19 and 20. The C-terminus of this fragment would be predicted to be within or close to a stretch of 11 amino acids with alternate Glu residues (731–741) followed by a proline-rich (45%) stretch of 20 amino acids. Antibody 89.1 reacted with the two fusion proteins derived from sequences within the 83K fragment region of the gene (Fig. 3c, tracks 3, 9), which locates the epitope for

Fig. 2 Nucleotide sequence of the P195 gene and the deduced amino-acid sequence of the open reading frame, obtained from the overlapping DNA clones shown in Fig. 1. Overlined, the putative signal peptide and membrane anchor sequences; ↓, the end of the putative signal peptide and the beginning of the amino-acid sequence for the 83K fragment (Fig. 4a); boxed region, the amino-acid repeat structure; *, potential N-glycosylation sites.

Fig. 3 Expression of specific fragments of the P195 gene as fusion proteins in *E. coli*. SDS-polyacrylamide gel electrophoresis (PAGE) analysis of total lysates from bacterial strains containing the plasmids pWRL 507 (tracks 1, 2), pME1 (tracks 3, 4), pME2 (tracks 5, 6), pME3 (tracks 7, 8) and pME4 (tracks 9, 10) grown in the presence of indole acrylic acid (+) or tryptophan (-). *a*, The proteins have been stained with Coomassie blue. *b*, *c*, The proteins have been transferred to nitrocellulose and probed with either a polyclonal antiserum against purified P195 (*b*) or monoclonal antibody 89.1 (*c*). The positions of standard *M_r* markers are shown. Plasmid constructions: pWRL507 (a gift from Dr M. Winther) is derived from pAT153 and contains the *trpE* control region and part of the coding sequence of the *trpE* gene product¹⁶ to the *Bgl*II site at nucleotide 966 followed by a synthetic 13-bp *Bgl*II-*Eco*RI adaptor. The recombinant plasmids contain: pME1, the 2.7-kb *Nde*I-*Eco*RI fragment from pPfg1 (nucleotides 876-3,559); pME2, the 2.4-kb *Eco*RI-*Hind*III fragment from pPfc1028 (nucleotides 3,555-5,920, with the *Hind*III site in the polylinker region); pME3, the 1.2-kb *Eco*RI-*Nde*I fragment from pPfc1028 (nucleotides 3,555-4,759); pME4, the 0.75-kb *Rsa*I-*Hind*III fragment from pPfc1013 (comprising nucleotides 1,065-1,816, the *Hind*III site is in the polylinker region).



this antibody between amino-acid residues 154 and 466.

Antisera raised against the partially purified fusion proteins as well as monoclonal antibodies were used to immunoprecipitate P195 and specific fragments from extracts of ³⁵S-methionine-labelled parasites. All the antibodies precipitated P195 (Fig. 4b). Antibody 89.1 (track 2) and antibodies against the fusion protein

produced by plasmid pME1 (track 6) precipitated the 83K fragment together with the intermediate processing fragments of about 150K and 110K that have been described previously⁶. The two monoclonal antibodies 111.2 and 111.4 did not bind to the 83K fragment or the intermediates (tracks 3 and 4). These two antibodies reacted with a different series of fragments, in particular, 30K and 46K species. The same polypeptides bound by 111.2 and 111.4, as well as the 150K species, were recognized by antisera raised against the fusion protein products of pME2 (track 7) and pME3 (track 8). From extracts of surface-labelled merozoites, antibody 89.1 immunoprecipitated only the 83K species⁸ while antibodies 111.2 and 111.4 precipitated both the 42K and 19K species, which appear to be complexed in a noncovalent association (data not shown). Since the 83K fragment has been located at the N-terminus of P195 by direct sequencing, these data support the tentative suggestion that the 42K merozoite surface fragment is derived from the C-terminus of P195. Antibody 89.1 does not detect the 83K fragment in newly invaded ring forms and the presence of this fragment in culture supernatants suggests that it is specifically shed during invasion. On the other hand, antibodies 111.2 and 111.4 react with newly invaded ring forms in the immunofluorescence test. This suggests that the fragment containing these specific epitopes is carried through into the red cell during invasion.

We thank Dr M. Winther for supplying the plasmid vector pWRL 507, Dr D. Walliker for supplying the Thai clones and Dr D. Stone for the amino-acid sequence analysis. R.T.S. is a recipient of a Heisenberg Fellowship from the Deutsche Forschungsgemeinschaft.

Received 3 May; accepted 24 July 1985.

- Holder, A. A. & Freeman, R. R. *Nature* **294**, 361-364 (1981).
- Holder, A. A., Freeman, R. R. & Newbold, C. I. *Molec. Biochem. Parasit.* **9**, 191-196 (1983).
- Holder, A. A. & Freeman, R. R. *Phil. Trans. R. Soc. B307*, 171-177 (1984).
- Hall, R. *et al. Nature* **311**, 379-382 (1984).
- Perrin, L. H. *et al. J. exp. Med.* **160**, 441-451 (1984).
- Holder, A. A. & Freeman, R. R. *J. exp. Med.* **156**, 1528-1538 (1982).
- Hall, R. *et al. Molec. Biochem. Parasit.* **11**, 61-80 (1984).
- Freeman, R. R. & Holder, A. A. *J. exp. Med.* **158**, 1647-1653 (1983).
- Holder, A. A. & Freeman, R. R. *J. exp. Med.* **160**, 624-629 (1984).
- Odink, K. G. *et al. FEBS Lett.* **173**, 108-112 (1984).
- Thaitong, S. *et al. Trans. R. Soc. Trop. Med. Hyg.* **78**, 242-245 (1984).
- Sanger, F., Nicklen, S. & Coulson, A. R. *Proc. natn. Acad. Sci. U.S.A.* **74**, 5463-5467 (1977).
- Maxam, A. M. & Gilbert, W. *Meih. Enzym.* **65**, 499-560 (1980).
- Staden, R. *Nucleic Acids Res.* **10**, 4731-4751 (1982).
- Staden, R. *Nucleic Acids Res.* **12**, 521-538 (1984).
- Nichols, B. P., van Cleemput, M. & Yanofsky, C. *J. molec. Biol.* **146**, 45-54 (1981).
- Baccanari, D. P., Tansik, R. L., Paterson, S. J. & Stone, D. *J. biol. Chem.* **259**, 12291-12298 (1984).
- Cheung, A. *et al. EMBOJ.* **4**, 1007-1011 (1985).

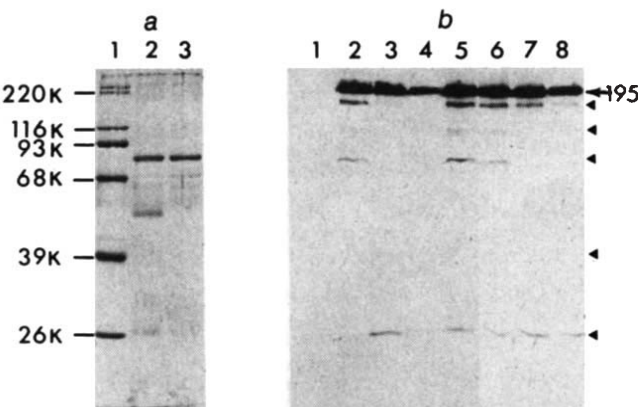


Fig. 4 Identification of the processed fragments within P195. *a*, Affinity purification of the 83K fragment from supernatants of *in vitro* cultures of *P. falciparum* for N-terminal sequencing¹⁷. Culture supernatants were passed through a column of antibody 89.1-Sepharose and after extensive washing the specifically bound material was eluted⁹ (track 2) and then passed through a protein A-Sepharose column to remove contaminant immunoglobulin (track 3). Polypeptides were analysed by SDS-PAGE on a 10% polyacrylamide gel and detected by Coomassie blue staining. In track 1 are standard *M_r*. *b*, SDS-PAGE analysis on a 10% polyacrylamide gel of polypeptides from extracts of ³⁵S-methionine-labelled *P. falciparum*-infected red cells. Extracts were immunoprecipitated with normal mouse serum (track 1), monoclonal antibody 89.1 (track 2), monoclonal antibody 111.2 (track 3), monoclonal antibody 111.4 (track 4), rabbit polyclonal anti-P195 (track 5), mouse antiserum raised against the fusion protein produced by the recombinant plasmids pME1 (track 6), pME2 (track 7) and pME3 (track 8). Monoclonal antibodies 111.2 and 111.4 were obtained using the procedure described previously⁶, except that the BALB/c mice were immunized intraperitoneally with 5×10^8 merozoites and then boosted 28 days later with a similar number of merozoites intravenously. Antisera to the fusion proteins produced in *E. coli* were prepared by immunizing mice intraperitoneally on two occasions 21 days apart with 100 μ g of protein in the presence of Freund's complete adjuvant. Antiserum was collected 7 days later.

Biological Sciences

The Amphibian Ear. By ERNEST GLEN WEVER. Princeton University Press:1985. Pp.488. ISBN 0-691-08365-7. £70.

Annals of the New York Academy of Sciences, Vol. 435. First Colloquium in Biological Sciences. WALTER N. SCOTT and FLEUR L. STRAND (eds). The New York Academy of Sciences:1984. Pp.621. ISBN 0-89766-264-4. \$140.

Basic Life Sciences, Vol. 31: Genetic Consequences of Nucleotide Pool Imbalance. FREDERICK J. DE SERRES (ed.). Plenum:1985. Pp.523. ISBN 0-306-41902-5. Np.

Behavioural Ecology: Ecological Consequences of Adaptive Behaviour. R.M. SIBLY and R.H. SMITH (eds). Blackwell Scientific:1985. Pp.620. Pbk ISBN 0-632-01413-X. Hbk £48; pbk £19.50.

Biological Membranes. D. CHAPMAN (ed.). Academic:1985. Pp.494. ISBN 0-12-168546-2. \$75, £55.

The Biology of Learning. P. MARLER and H.S. TERRACE (eds). Springer-Verlag:1985. Pp.738. ISBN 0-387-13923-0/3-540-13923-0. DM 88.

Blue Light Effects in Biological Systems. HORST SENER (ed.). Springer-Verlag:1985. Pp.538. ISBN 0-387-13462-X/3-540-13462-X. DM 142.

The Cell Division Cycle in Plants. J.A. BRYANT and D. FRANCIS (eds). Cambridge University Press:1985. Pp.258. ISBN 0-521-30046-0. £18.50, \$32.50.

Cell and Muscle Motility, Vol. 6. JERRY W. SHAY (ed.). Plenum:1985. Pp.377. ISBN 0-306-41785-5. Np.

Concepts in Marine Pollution: Measurements. HARRIS H. WHITE (ed.). Maryland Sea Grant:1985. Pp.743. ISBN 0-943676-18-5. \$12.50.

Context and Learning. PETER D. BALSAM and ARTHUR TOMIE (eds). Erlbaum:1985. Pp.419. ISBN 0-89859-442-1. £39.95.

Controlling Events in Meiosis. CLIVE W. EVANS and H.G. DICKINSON (eds). The Biochemical Society:1985. Pp.394. ISBN 0-9508709-1-9. £32, \$60.

Coordination of Motor Behaviour. B.M.H. BUSH and F. CLARAC (eds). Cambridge University Press:1985. Pp.324. ISBN 0-521-26425-1. £20, \$34.50.

The Developing Child, 4th Edn. By HELEN BEE. Harper & Row:1985. Pp.582. ISBN 0-06-350131-7. £10.50.

Development, Organization, and Processing in Somatosensory Pathways: Neurology and Neurobiology, Vol. 14. MARK ROWE and WILLIAM D. WILLIS Jr (eds). Alan R. Liss:1985. Pp.411. ISBN 0-8451-2716-0. £63.

DNA Methylation: Biochemistry and Biological Significance. AHARON RAZIN, HOWARD CEDAR and ARTHUR D. RIGGS (eds). Springer-Verlag:1984. Pp.392. ISBN 3-540-96038-4/0-387-96038-4. DM 188.

Endocrine Tumours: The Pathobiology of Regulatory Peptide-Producing Tumours. By J.M. POLAK and S.R. BLOOM. Churchill Livingstone:1985. Pp.288. ISBN 0-443-02979-2. £35.

Environmental Physiology and Biochemistry of Insects. KLAUS H. HOFFMANN (ed.). Springer-Verlag:1985. Pp.296. ISBN 3-540-13762-9/0-387-13762-9. DM 98, \$34.40.

Environmental Problems/Behavioral Solutions. By JOHN D. CONE and STEVEN C. HAYES. Cambridge University Press:1985. Pp.284. Pbk ISBN 0-521-31973-0. Pbk £6.95.

The Enzymes of Biological Membranes, 2nd Edn. Vol. 3. Membrane Transport. ANTHONY N. MARTONOSI (ed.). Plenum:1985. Pp.676. ISBN 0-306-41453-8.

Enzymes, Receptors and Carriers of Biological Membranes: A Laboratory Manual. A. AZZI, U. BRODBECK and P. ZAHLER (eds). Springer-Verlag:1984. Pp.165. Pbk ISBN 0-387-13751-3/3-540-13751-3. Pbk DM 29.

Essays in Biochemistry, Vol. 19, 1983. P.N. CAMPBELL and R.D. MARSHALL (eds). Academic Press:1983. Pp.208. ISBN 0-12-158119-5. \$18, £10.80.

Evolutionary Ecology of Marsupials. By ANTHONY K. LEE and ANDREW COCKBURN. Cambridge University Press:1985. Pp.274. ISBN 0-521-25292-X. £30, \$54.50.

Handbook of Neurochemistry, 2nd Edn. ABEL LAJTHA (ed.). Plenum:1985. Pp.792. ISBN 0-306-41744-8. Np.

Hand Function and the Neocortex: Experimental Brain Research Supplementum 10. A.W. GOODWIN and I. DARIAN-SMITH (eds). Springer-Verlag:1985. Pp.314. ISBN 3-540-13948-6. DM 98.

Hematopoietic Stem Cells. DAVID W. GOLDE and FUMIMARO TAKAKU (eds). Dekker:1985. Pp.379. ISBN 0-8247-7224-5. \$69.75.

Higher Order Structure in the Nucleus. P.R. COOK and R.A. LASKEY (eds). The Biochemical Society:1985. Pp.234. ISBN 0-9508709-4-3. £10, \$19.

How Animals See: Other Visions of Our World. By SANDRA SINCLAIR. Croom Helm:1985. Pp.146. ISBN 0-7099-3336-3. £12.95.

Inclusion Compounds, Vol. 3. J.L. ATWOOD, J.E.D. DAVIES and D.D. MACNICOL (eds). Academic:1985. Pp.669. ISBN 0-12-067103-4. \$98, £65.

Intestinal Microbiology. By B.S. DRASAR and P.A. BARROW. Van Nostrand Reinhold:1985. Pp.80. Pbk ISBN 0-442-30598-2. Pbk £5.25.

The Lion-Tailed Macaque: Status and Conservation. Monographs in Primatology, Vol. 7. PAUL G. HELTNE (ed.). Alan R. Liss:1985. Pp.411. ISBN 0-8451-3406-X. £44.

Mathematical Modelling of Dynamic Biological Systems, 2nd Edn. By LUDWIG FINKELSTEIN and EWART R. CARSON. Research Studies Press:1985. Pp.355. ISBN 0-86380-024-6. £43.50.

Mediterranean Marine Ecosystems. MARIA MORAITOU-APOSTOLOPOULOU and VASSILI KIORTSIS (eds). Plenum:1985. Pp.407. ISBN 0-306-41910-6. \$65.

Methods in Enzymology. LUTZ BIRNBAUMER and BERT W. O'MALLEY (eds). Academic:1985. Pp.891. ISBN 0-12-182009-2. \$85, £65.

Microbial Extrachromosomal Genetics. By STEPHEN OLIVER and TERENCE BROWN. Van Nostrand Reinhold:1985. Pp.85. Pbk ISBN 0-442-30587-7. Pbk £5.25.

Microfungi on Land Plants: An Identification Handbook. By MARTIN B. ELLIS and J. PAMELA ELLIS. Croom Helm:1985. Pp.818. ISBN 0-7099-0950-0. £45.

Microorganisms as Model Systems for Studying Evolution. ROBERT P. MORTLOCK (ed.). Plenum:1985. Pp.326. ISBN 0-306-41788-X. Np.

Modern Cellular Automata: Theory and Applications. By KENDALL PRESTON and MICHAEL J.B. DUFF. Plenum:1985. Pp.340. ISBN 0-306-41737-5. Np.

Modern Chromatographic Analysis of the Vitamins. ANDRÉ P. DE LEENHEER, WILLY E. LAMBERT and MARCEL G.M. DE RUYTER (eds). Dekker:1985. Pp.576. ISBN 0-8247-7221-0. \$102, SwFr. 255.

Monkeys and Apes. LONGMAN NATURE LIBRARY. Longman:1985. Pp.127. ISBN 0-582-89214-7. £3.95.

Neural Grafting in the Mammalian CNS. ANDERS BJÖRKLUND and ULF STENEVI (eds). Elsevier:1985. Pp.709. ISBN 0-444-80622-9. \$157.50, Dfl. 425.

Nutrition of Plantation Forests. G.D. BOWEN and E.K.S. NAMBIAR (eds). Academic:1985. Pp.516. ISBN 0-12-120980-6. \$75, £49.50.

Owls and Birds of Prey. LONGMAN NATURE LIBRARY. Longman:1985. Pp.128. ISBN 0-582-89215-5. £3.95.

Pediatrics Update. ARTHUR H. MOSS. Elsevier:1985. Pp.412. ISBN 0-444-00922-1. \$59.50.

Plant-Microbe Interactions: Molecular and Genetic Perspectives. TSUNE KOSUGE and EUGENE W. NESTER (eds). Collier Macmillan:1985. Pp.444. ISBN 0-02-949470-2. £42.

The Reticuloendothelial System: A Comprehensive Treatise. JOHN W. HADDEN and ANDOR SZENTIVANYI (eds). Plenum:1985. Pp.469. ISBN 0-306-41792-8. Np.

Siderophores from Microorganisms and Plants: Structure and Bonding, Vol. 58. M.J. CLARKE et al. (eds). Springer-Verlag:1984. Pp.143. ISBN 3-540-13649-5/0-387-13649-5. DM 86.

Smoking and the Lung. G. CUMMING and G. BONSIGNORE (eds). Plenum:1984. Pp.504. ISBN 0-306-41828-2. Np.

Successful Nature Watching: A Step-by-Step Guide to Watching Wildlife. By DEREK HALL, ANDREW CLEAVE and PAUL STERRY. Hamlyn:1985. Pp.240. ISBN 0-600-30602-X. £7.95.

A Theory of Forest Dynamics: The Ecological Implications of Forest Succession Models. By HERMAN H. SHUGART. Springer-Verlag:1985. Pp.278. ISBN 0-387-96000-7/3-540-96000-7. DM 89.

Transaminases. PHILIPP CHRISTEN and DAVID E. METZLER. Wiley:1985. Pp.643. ISBN 0-471-08501-4. £158.

Trichinellosis. CHARLES W. KIM (ed.). The State University of New York Press:1985. Pp.343. ISBN 88706-091-9. \$49.50.

The Virulence of Escherichia coli: Review and Methods. M. SUSSMAN (ed.). Academic:1985. Pp.473. ISBN 0-12-677520-6. \$75, £60.

Viruses and Cancer. P.W.J. RIGBY and N.M. WILKIE (eds). Cambridge University Press:1985. Pp.323. ISBN 0-521-26867-2. \$64.50.

Zoophysiology, Vol. 17: The Control of Fish Migration. By R.J.F. SMITH. Springer-Verlag:1985. Pp.243. ISBN 3-540-13707-6/0-387-13707-6. DM 136, \$47.80.

Applied Biological Sciences

Advances in Perinatal Medicine, Vol. 4. AUBREY MILUNSKY, EMANUEL A. FRIEDMAN and LOUIS GLUCK (eds). Plenum:1985. Pp.329. ISBN 0-306-41840-1. Np.

Anaerobic Treatment of Forest Industry Wastewaters. P. RANTALA and A. LUONSI (eds). Pergamon:1985. Pp.326. Pbk ISBN 0-08-032729-X. Pbk £27.50, \$44.

Asbestiform Fibers: Nonoccupational Health Risks. COMMITTEE ON NONOCCUPATIONAL RISKS OF ASBESTIFORM FIBERS. National Academy Press:1984. Pp.334. Pbk ISBN 0-309-03446-9. Pbk £25.85.

General

Absorption and the Gibbs Surface Excess. By D.K. CHATTORAJ and K.S. BIRDI. Plenum:1984. Pp.471. ISBN 0-306-41334-5. \$59.50.

Abstract Measurement Theory. By LOUIS NARENS. Massachusetts Institute of Technology:1985. Pp.334. ISBN 0-262-14037-3. Np.

Abstracts and Indexes in Science and Technology: A Descriptive Guide, 2nd Edn. By DOLORES B. OWEN. Scarecrow Press, 52 Liberty St, PO Box 656, Metuchen, NJ 08840, USA:1985. Pp.235. ISBN 0-8108-1712-8. \$17.50.

The Amazon. HARALD SIOLI (ed.). Dr W. Junk:1984. Pp.763. ISBN 90-6193-108-8. \$172, £114.50.

Anticipated and Abnormal Plant Transients in Light Water Reactors, Vols 1 & 2. PAMELA L. LAS-SAHN, DEBU MAJUMDAR and GEORGE F. BROCKETT (eds). Plenum:1984. Vol. 1 ISBN 0-306-41718-9. Vol. 2 ISBN 0-306-41718-9. £225.

Assessing Sex Bias in Testing: A Review of the Issues and Evaluations of 74 Psychological and Educational Tests. By PAULA SELKOW. Greenwood Press, Westport, Connecticut/Westport Publications, 3 Henrietta Street, London:1984. Pp.264. £27.95.

Biotechnology Engineers: Biographical Directory. By OSKAR R. ZABORSKY and DONNA K. ZUBRIS. OMEC:1985. Pp.360. Pbk ISBN 0-93128-05-1. Pbk \$29.95.

Buried Structures: Static and Dynamic Strength. By P.S. BULSON. Chapman and Hall:1985. Pp.227. ISBN 0-412-21560-8. £25, \$42.50.

Ceramic Theory and Cultural Process. By DEAN E. ARNOLD. Cambridge University Press:1985. Pp.268. ISBN 0-521-25262-8. £19.50.

The Challenge of Smallholding. By SEDLEY SWEENEY. Oxford University Press:1985. Pp.222. Pbk ISBN 0-19-286018-6. Pbk £3.95.

Chaos. By HAO BAI-LIN. World Scientific:1984. Pp.576. Pbk ISBN 9971-966-50-6. Pbk £25.30.

Chemical Processes in Lakes. WERNER STUMM (ed.). Wiley:1985. Pp.435. ISBN 0-471-88261-5. £69.35.

The Christian Frame of Mind. By THOMAS F. TORRANCE. Handzel Press, 33 Montgomery St, Edinburgh EH7 5JX:1985. Pp.62. Pbk ISBN 0-905312-43-0. Pbk £3.50.

Cry of the Kalahari. By MARK and DELIA OWENS. Harvill:1985. Pp.341. ISBN 0-00-216357-8. £9.95.